

Figure 1

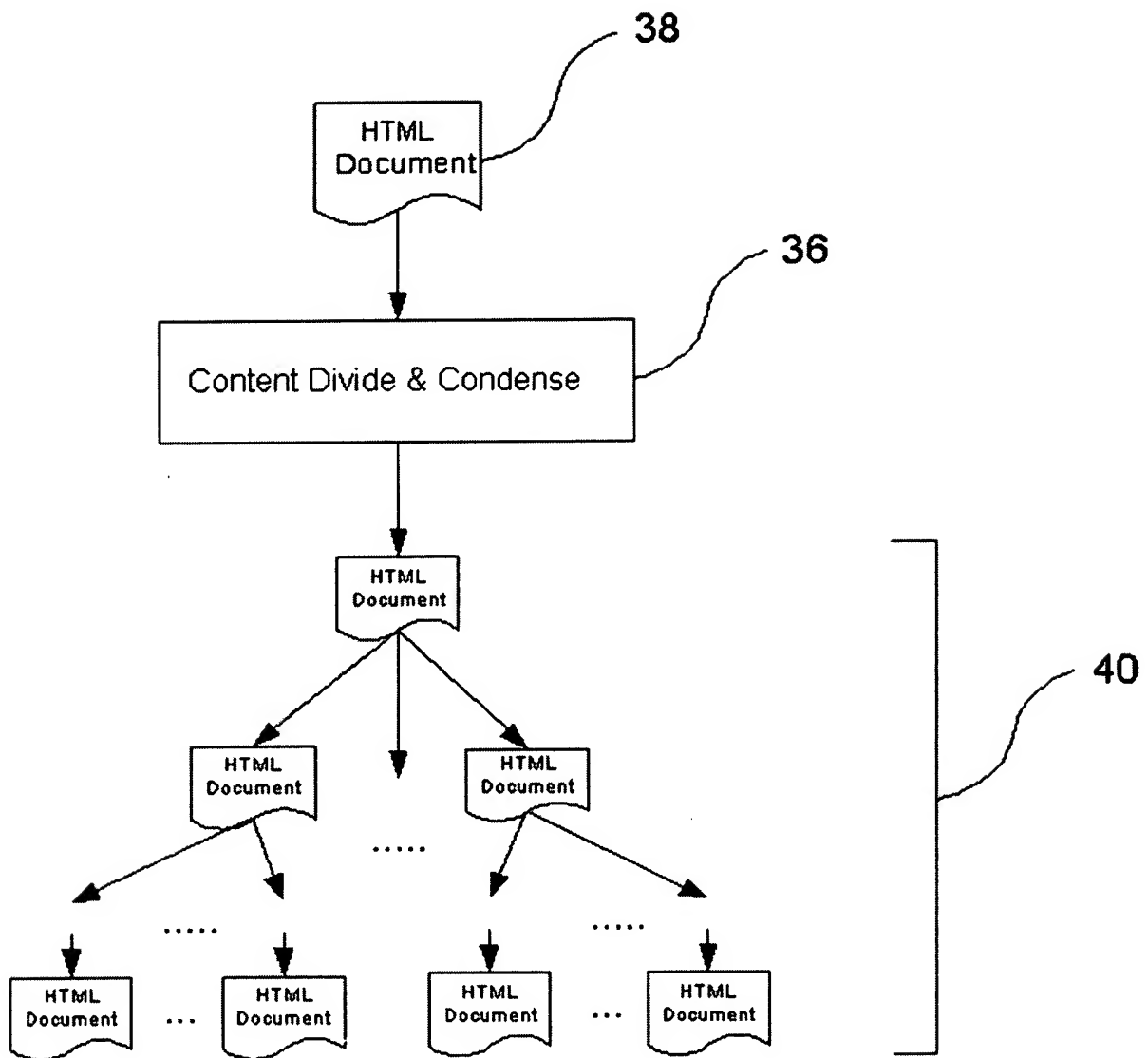


Figure 2

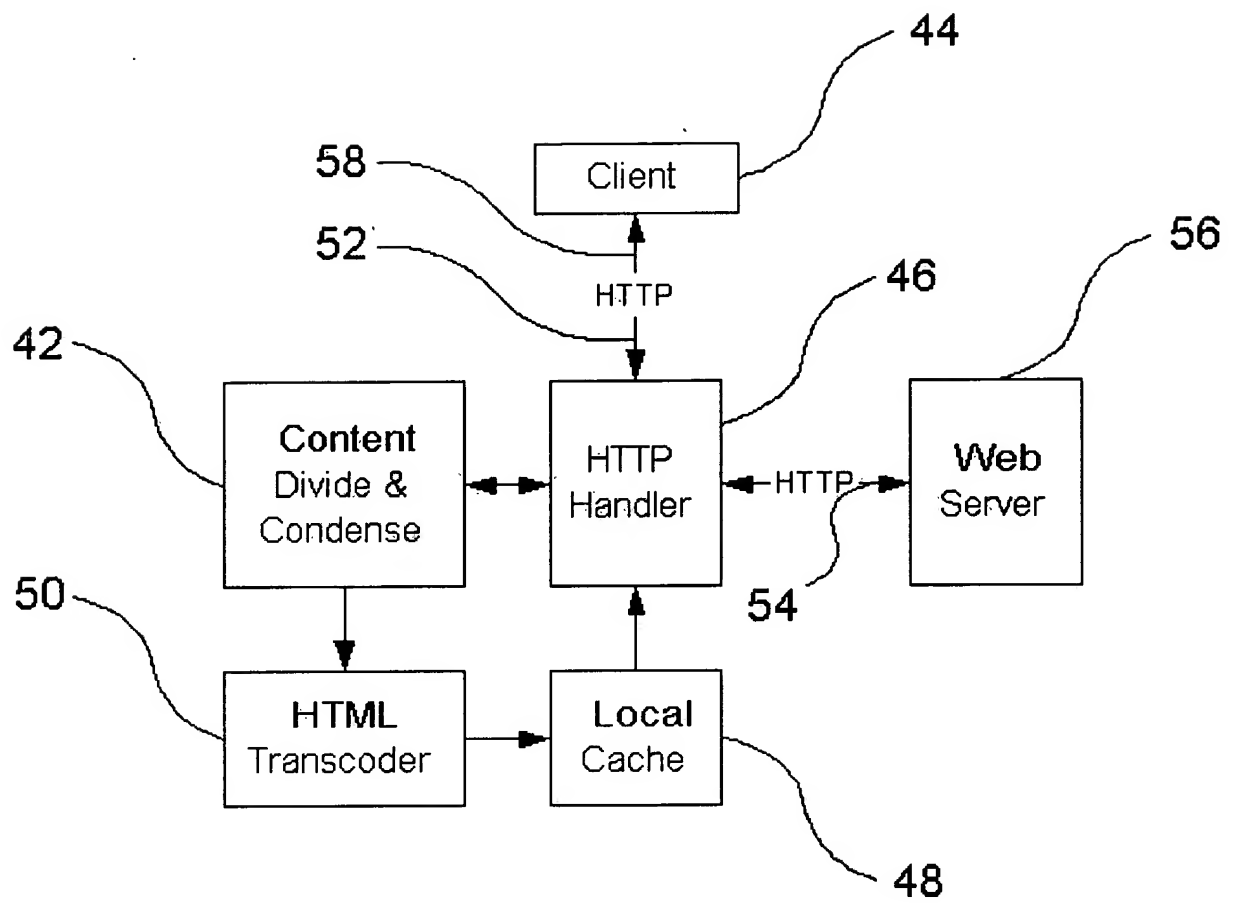


Figure 3

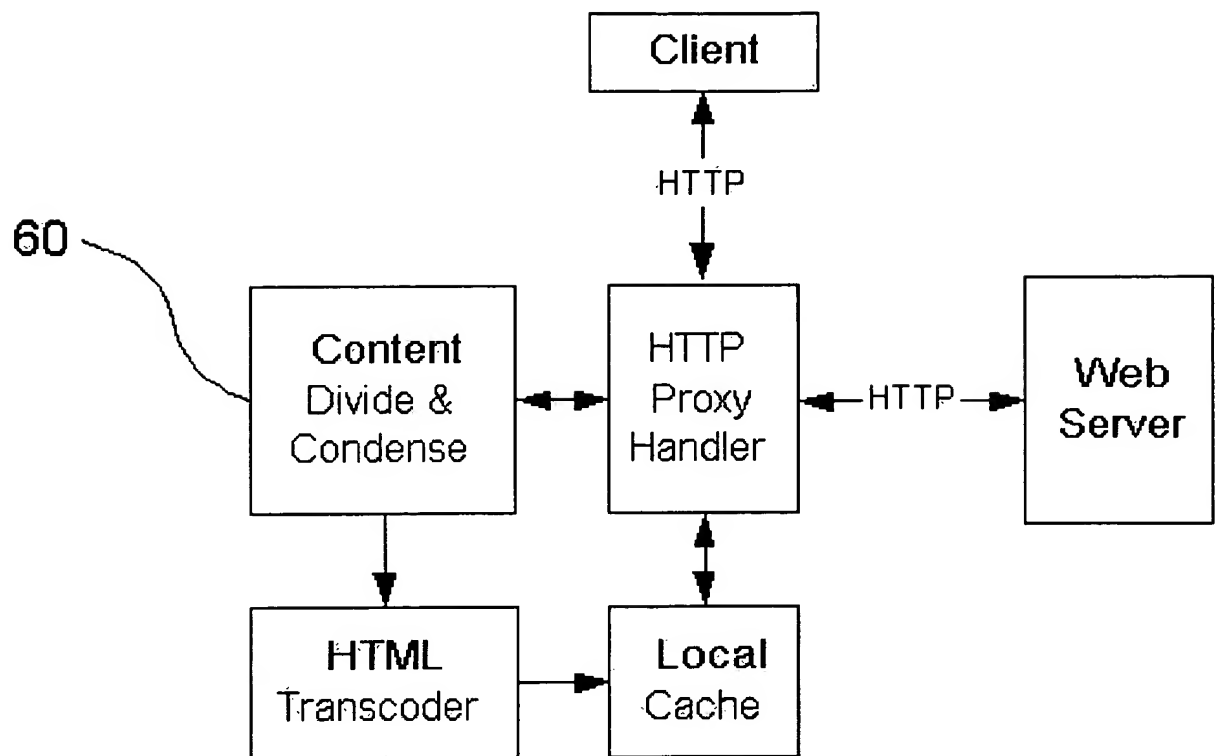


Figure 4

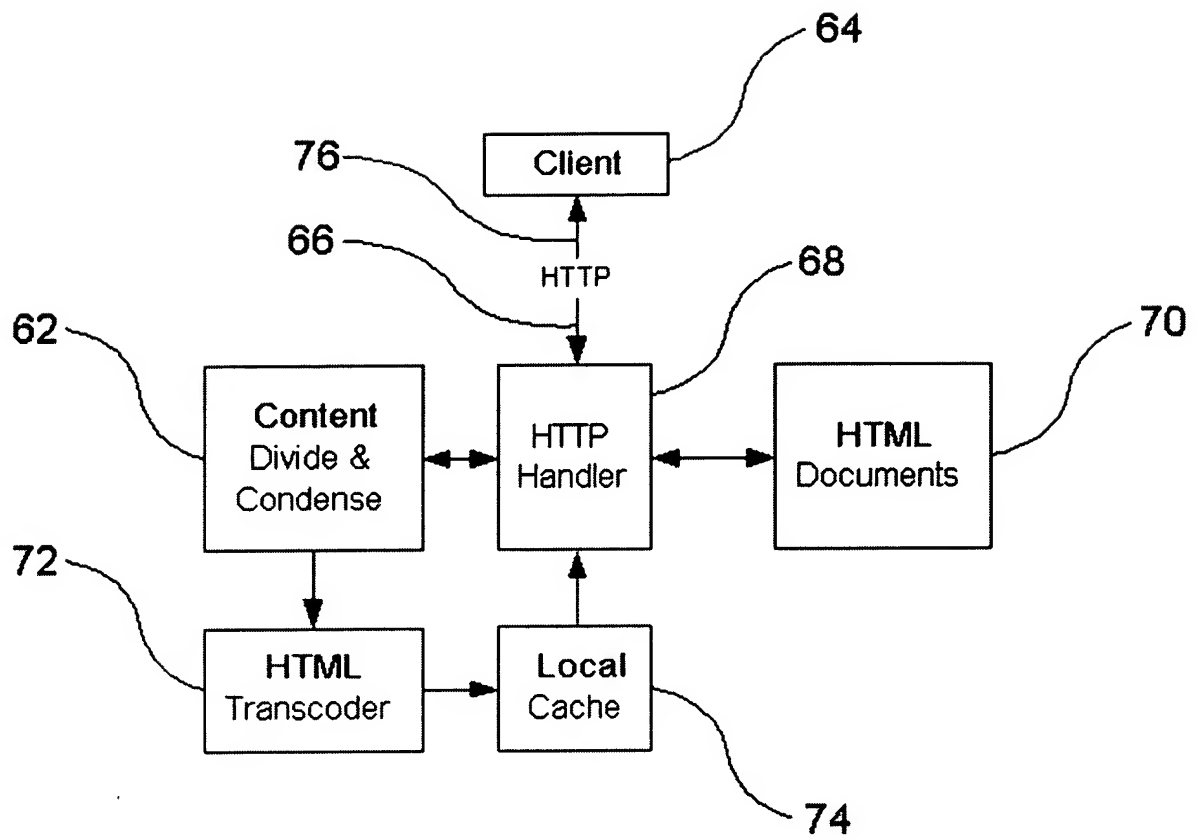


Figure 5

```

<html><head><title> Example
</title>
</head>
<body bgcolor="blue"
text="yellow"><center><h2>Colorful page with
blue and yellow.</h2></center> </body>
</html>

```

77a

Parsing

77b

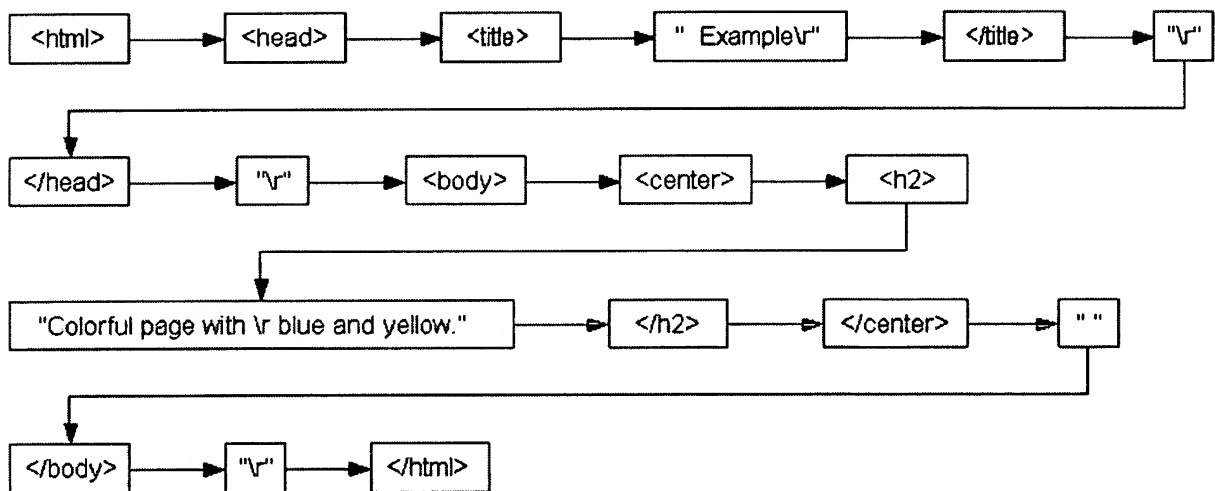


Figure 6

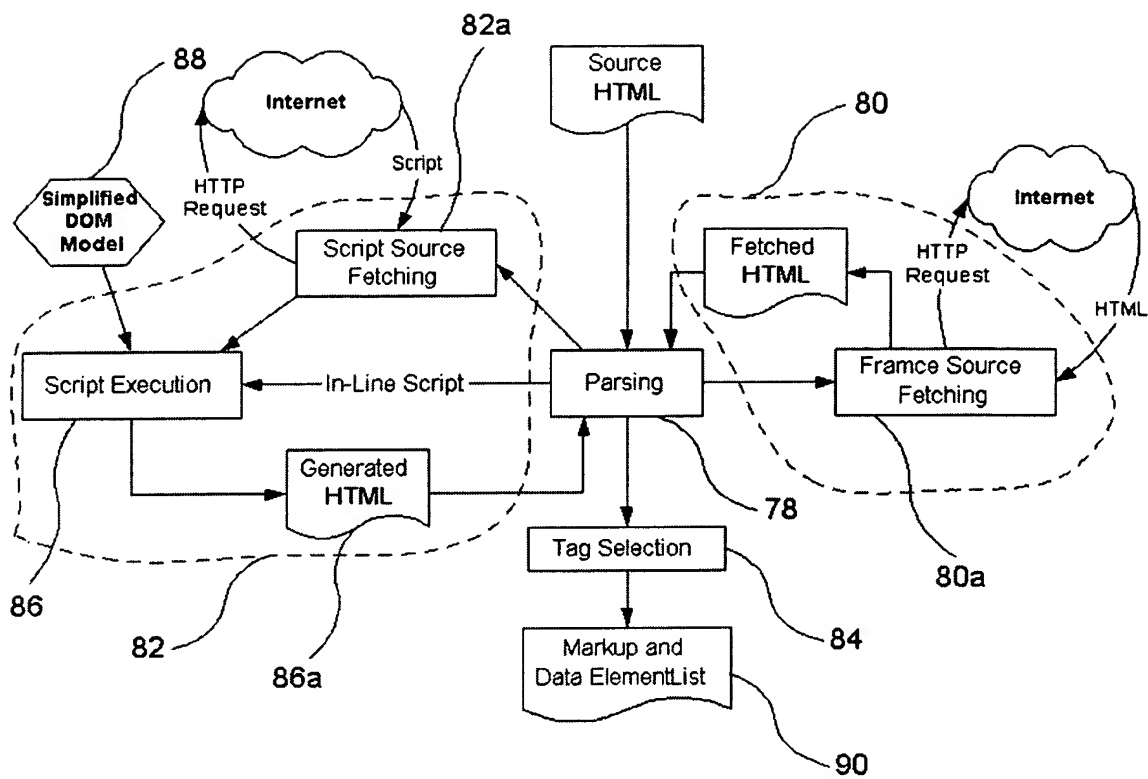


Figure 7

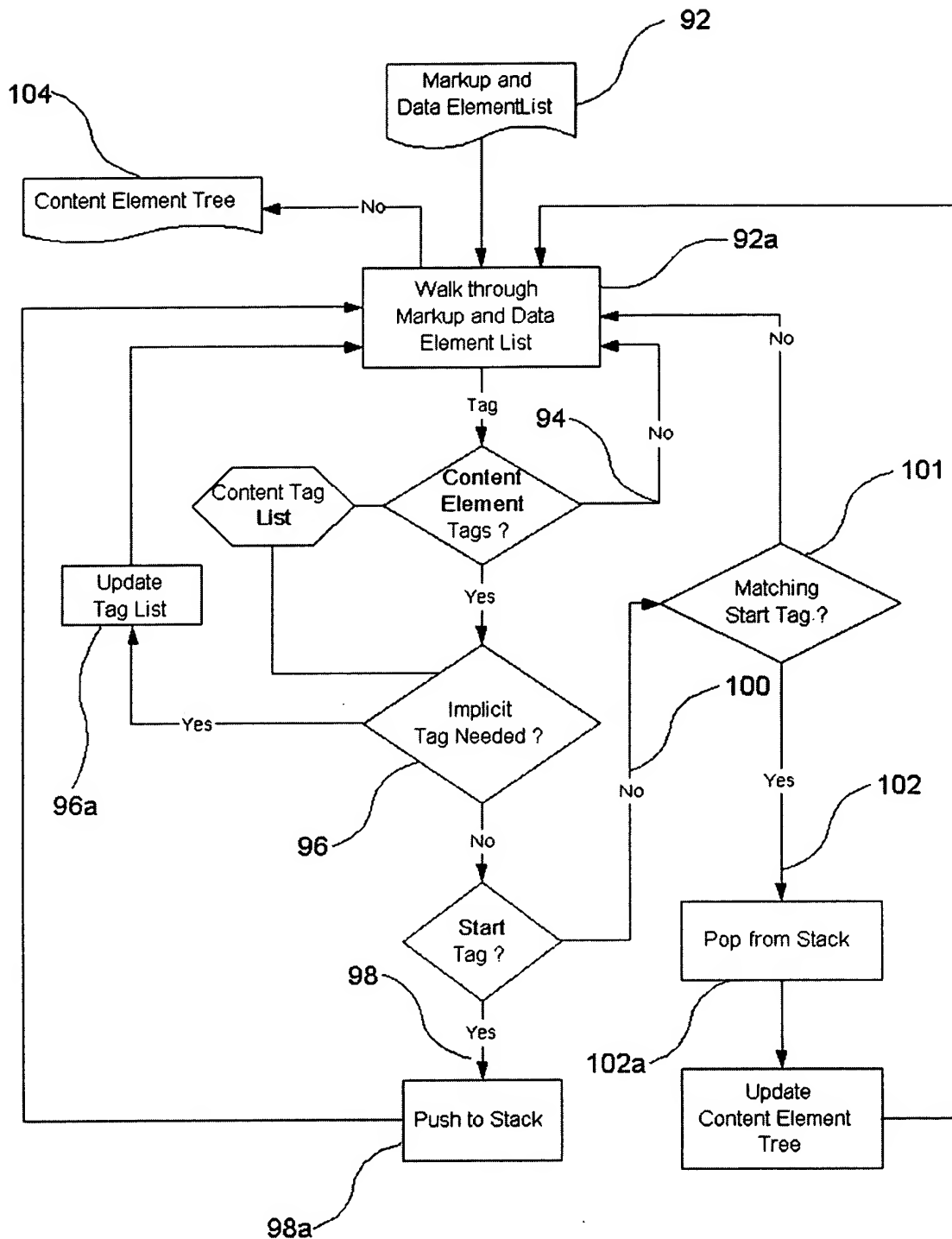


Figure 8



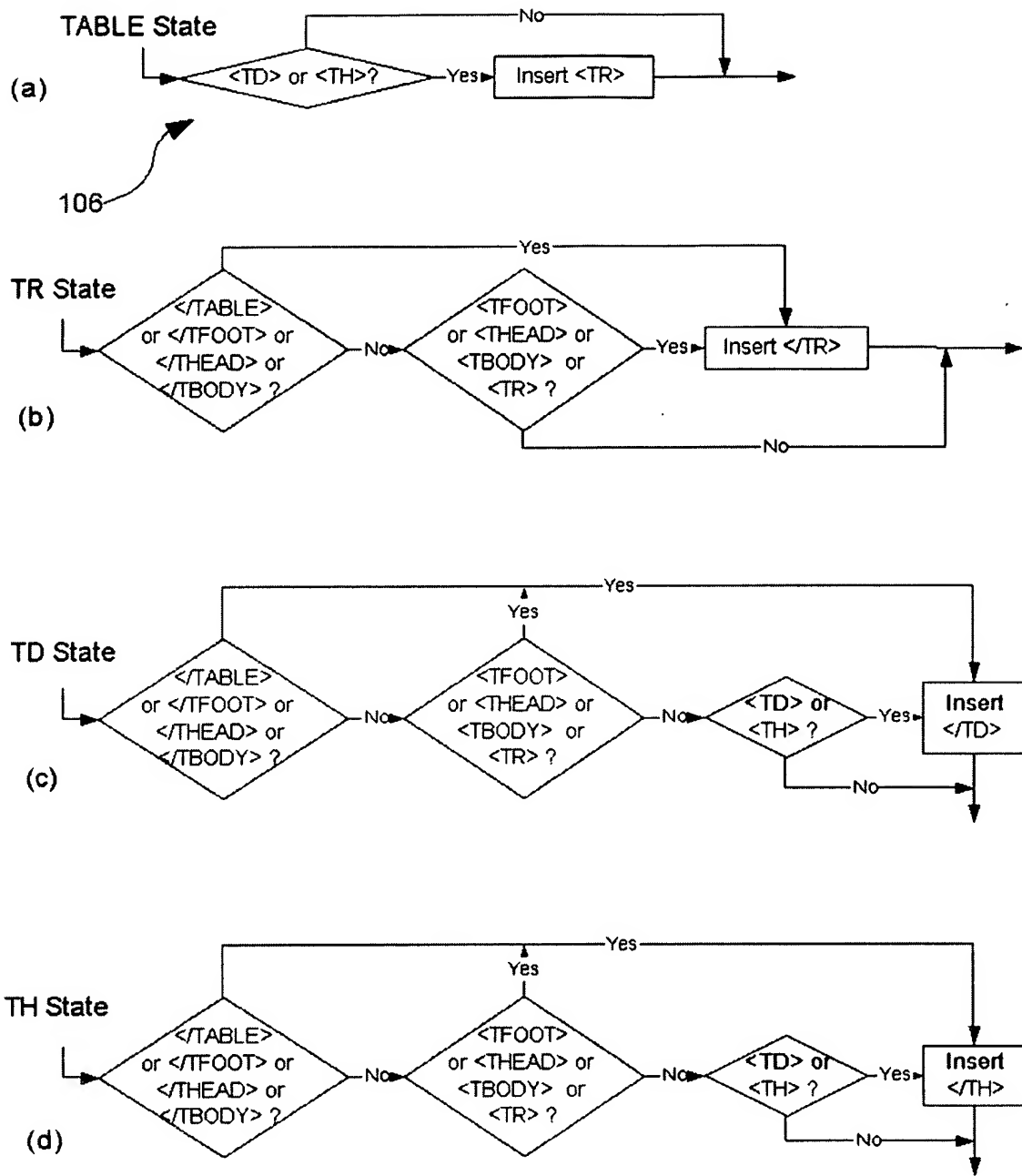


Figure 9a

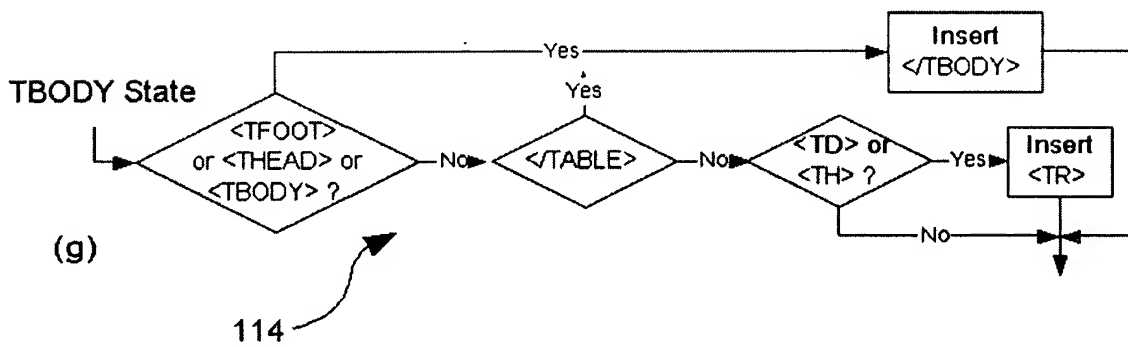
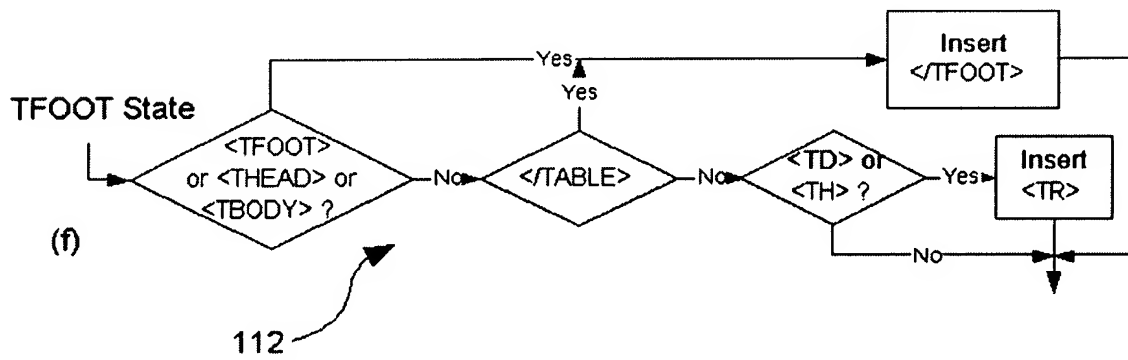
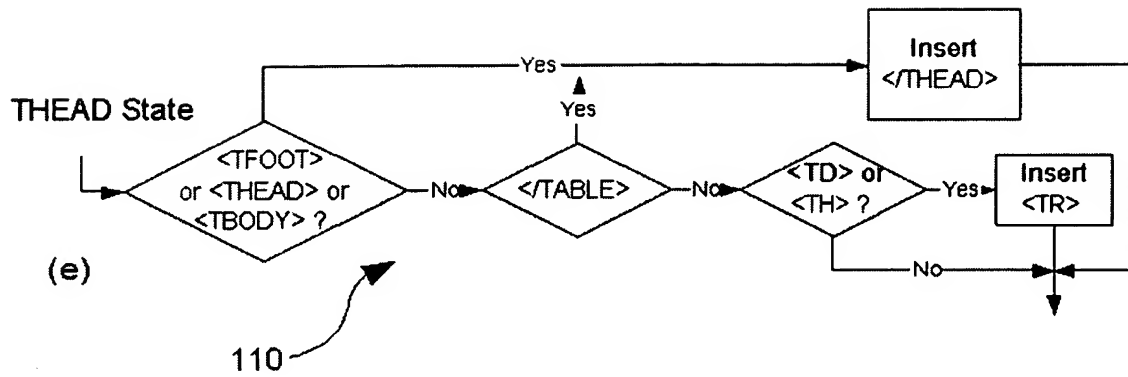


Figure 9b

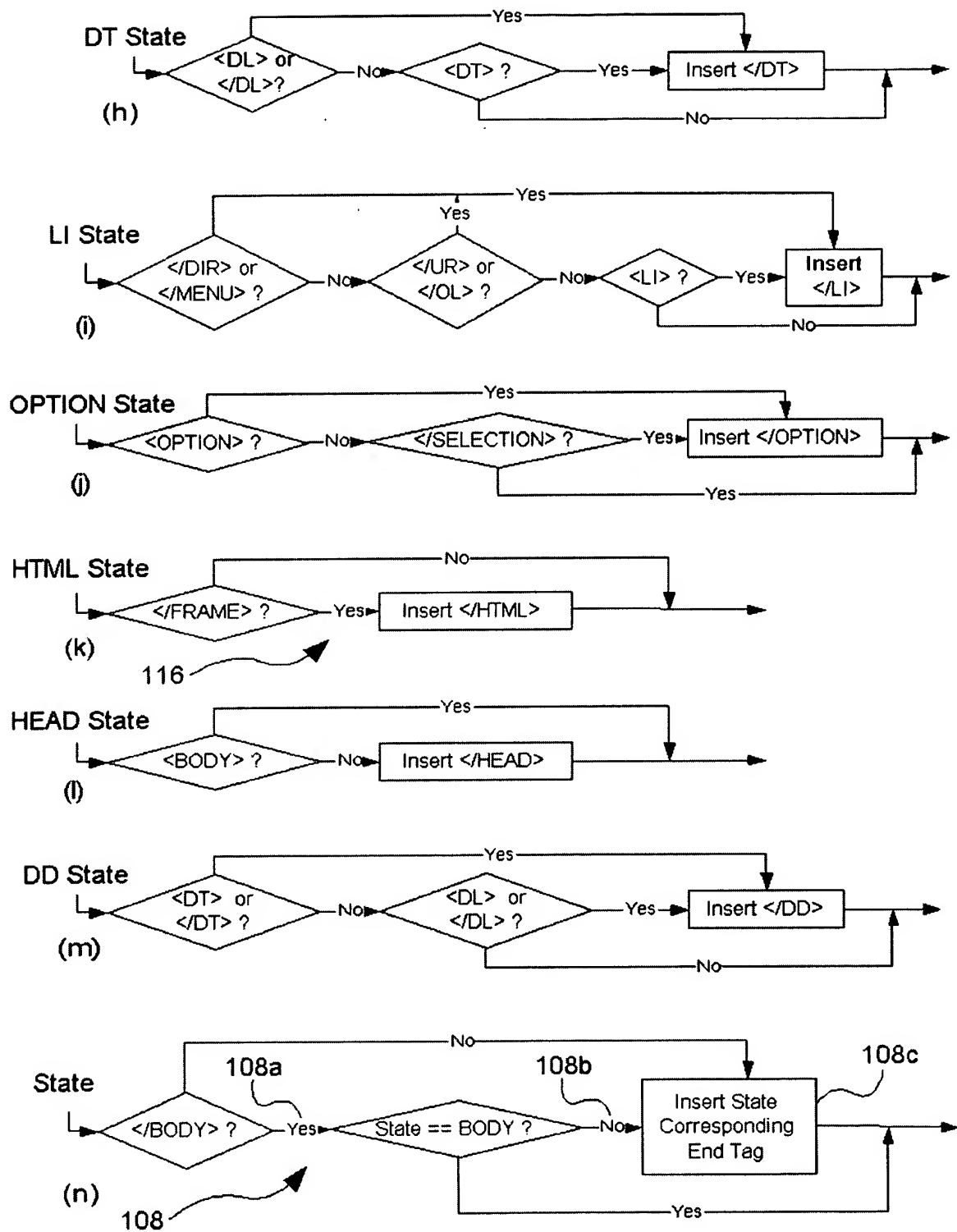


Figure 10

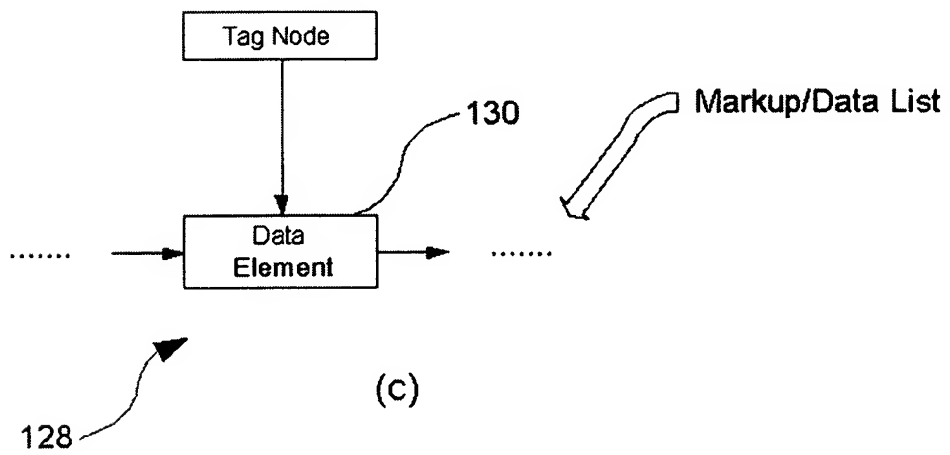
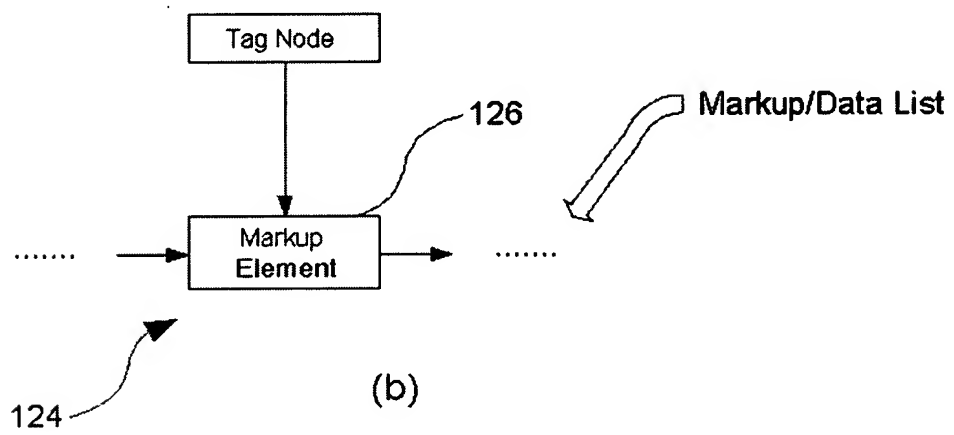
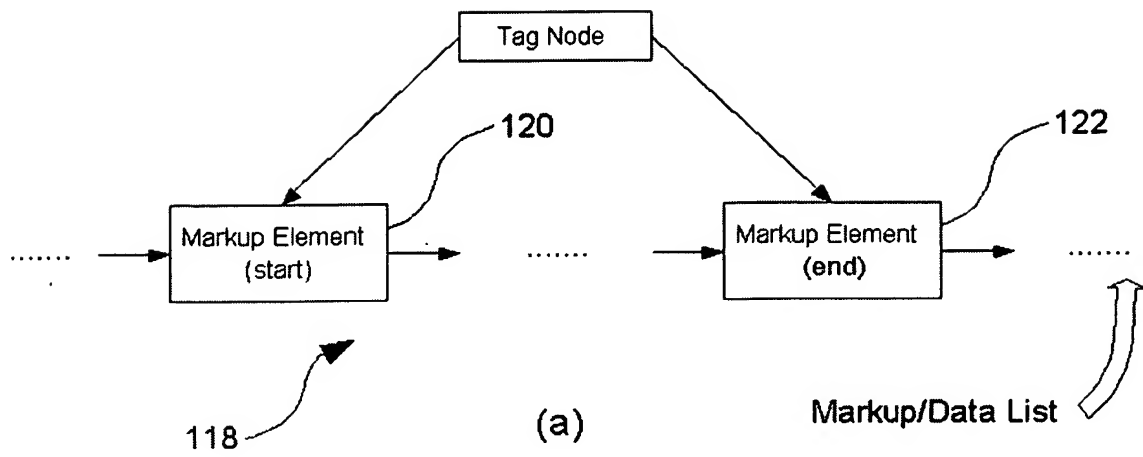


Figure 11

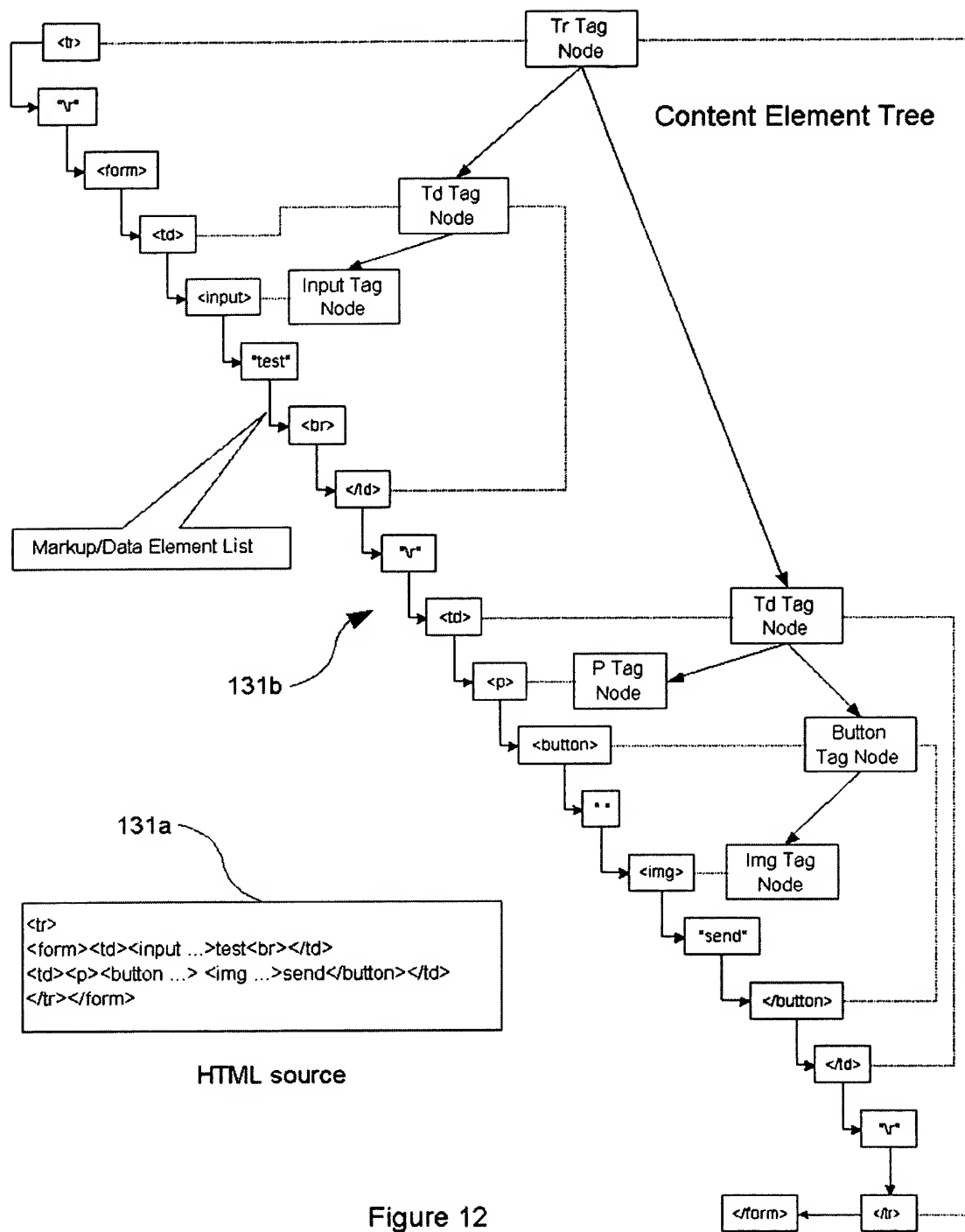


Figure 12

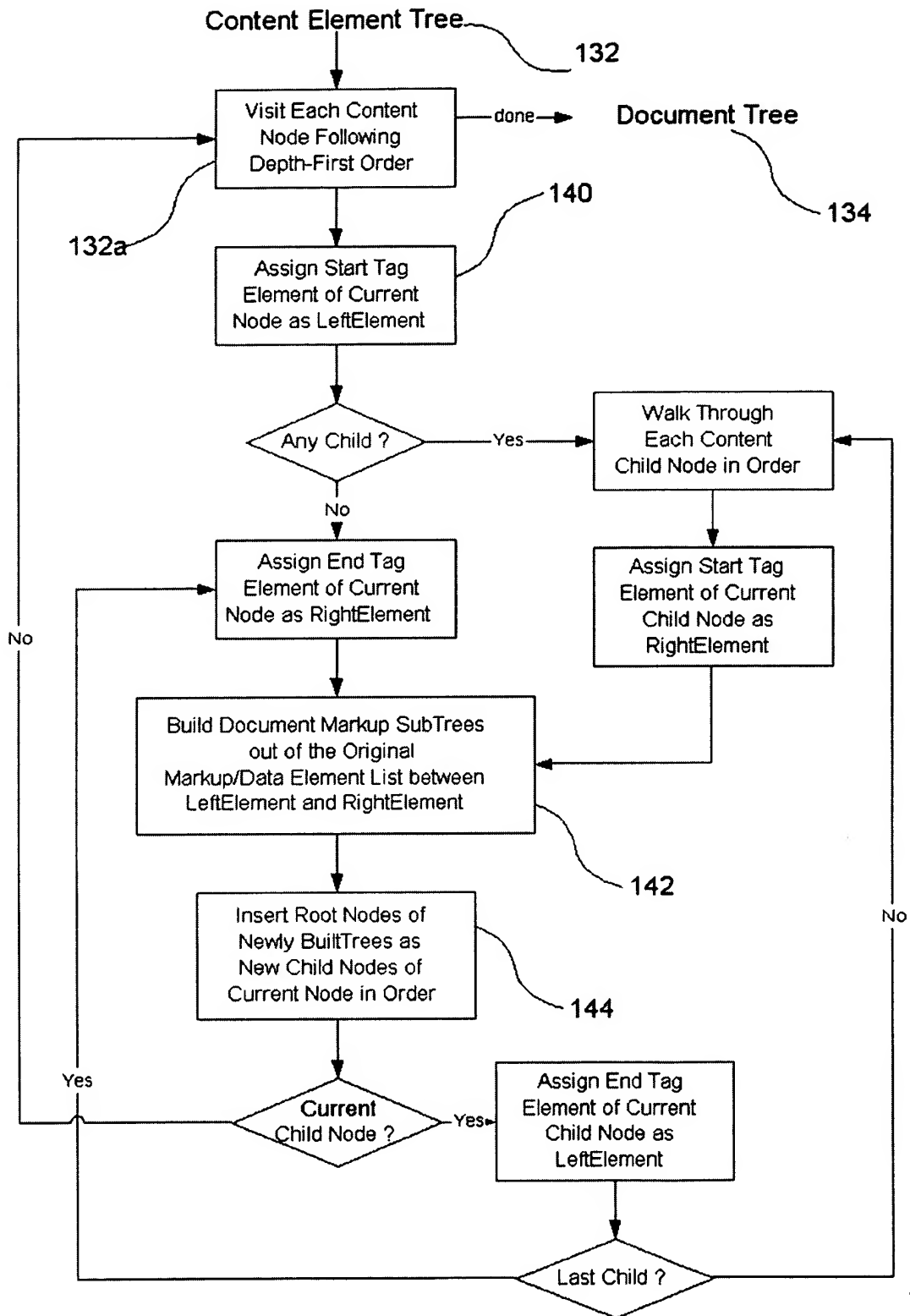


Figure 13

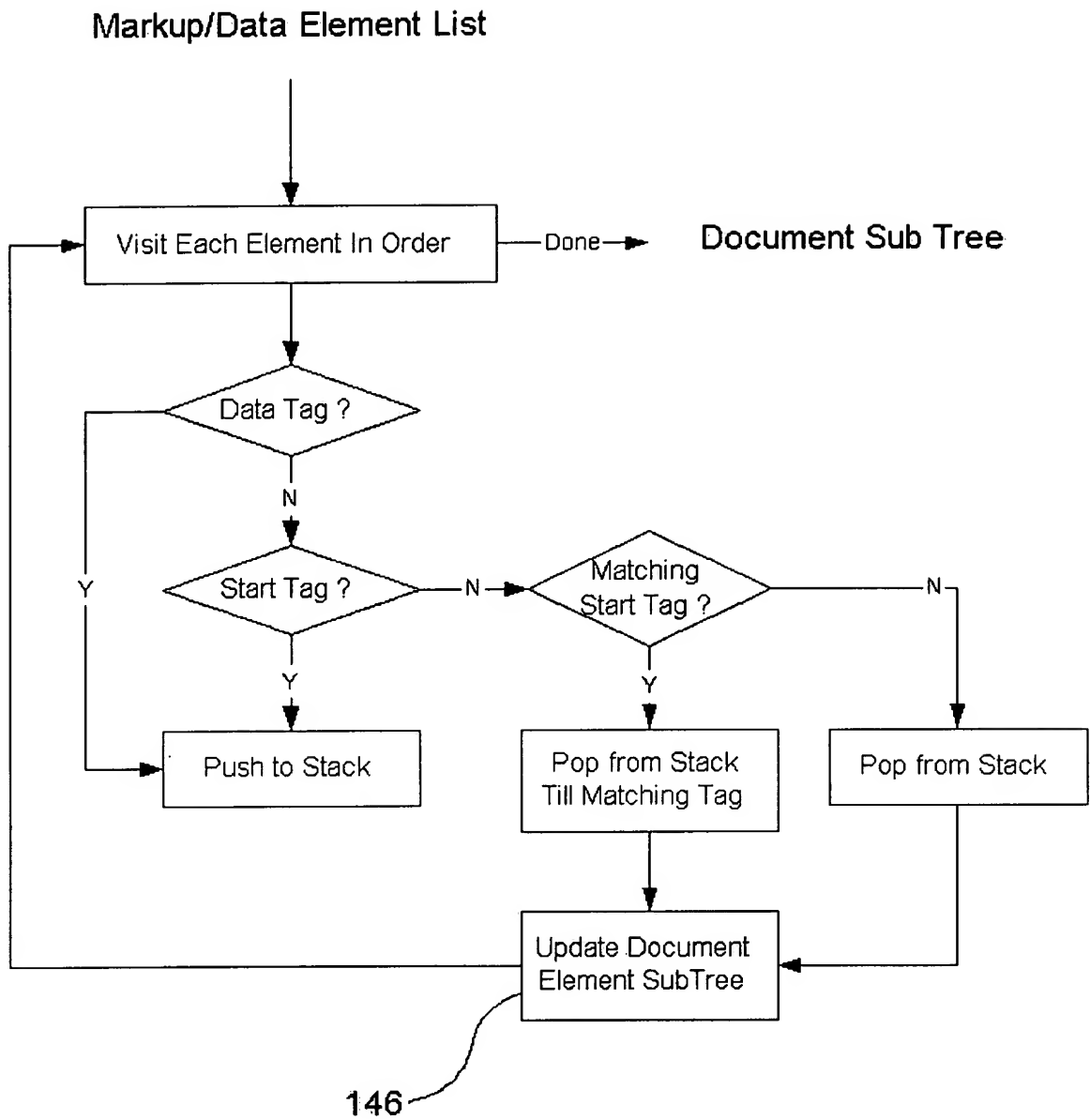


Figure 14

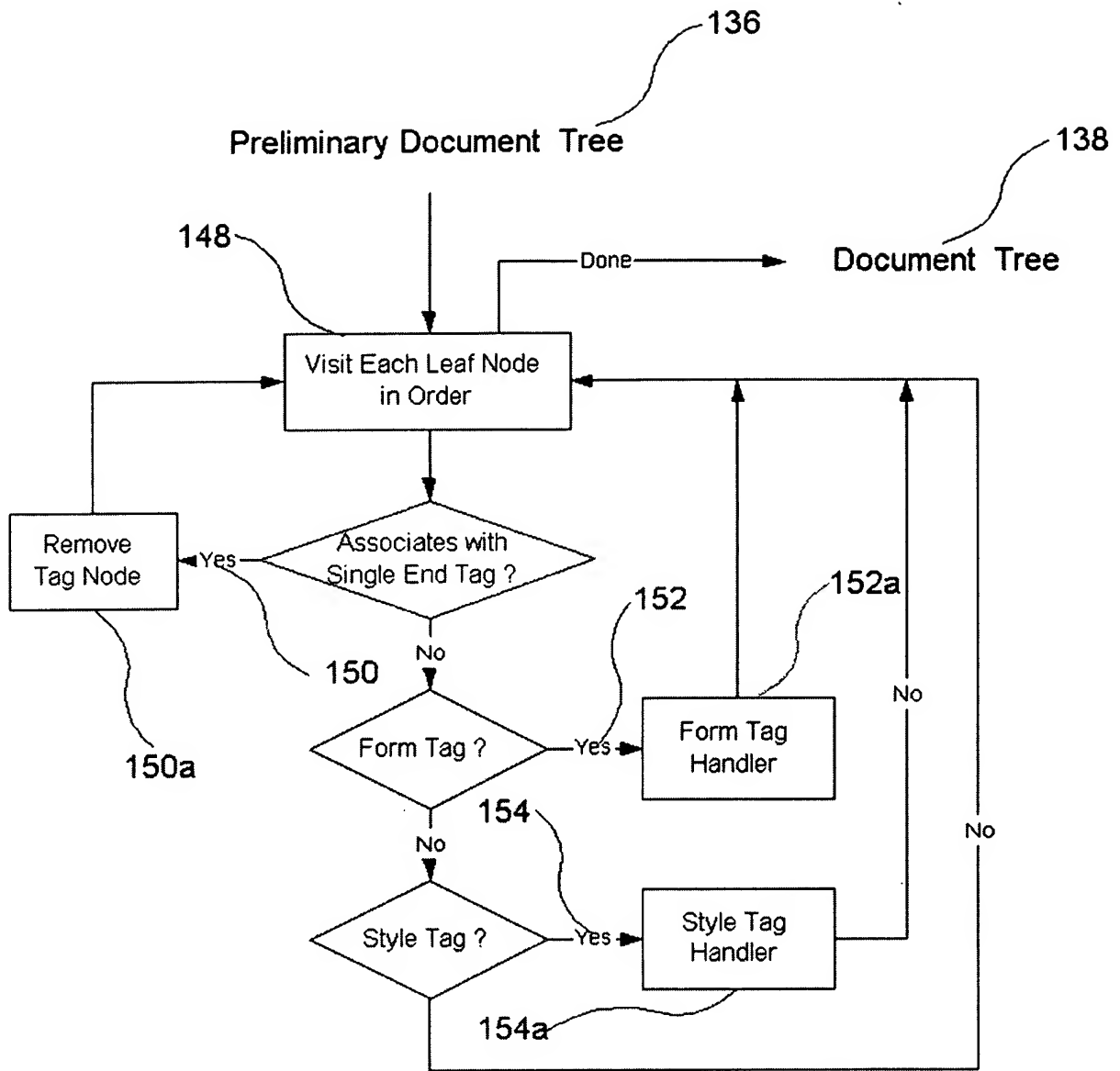


Figure 15



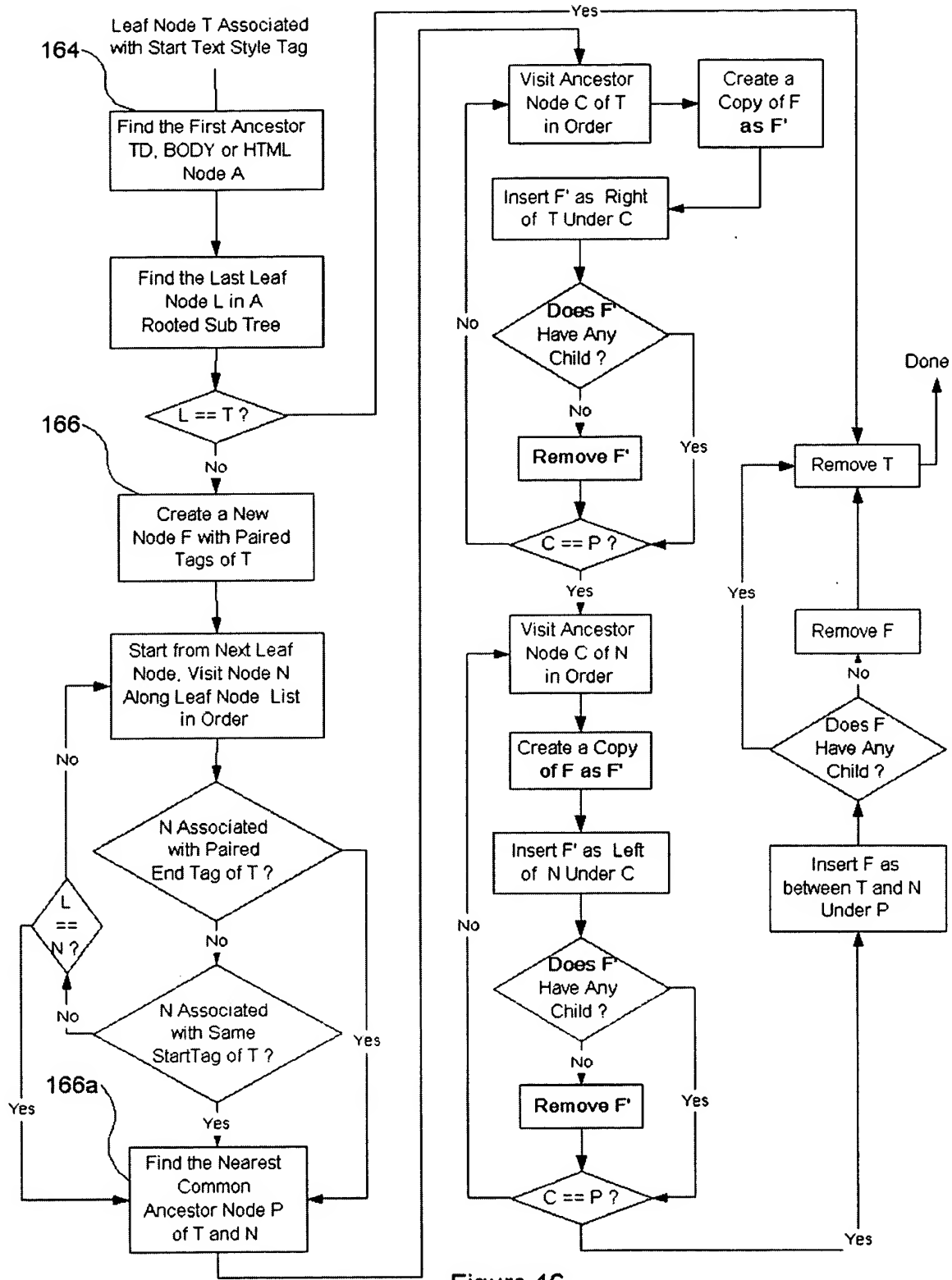
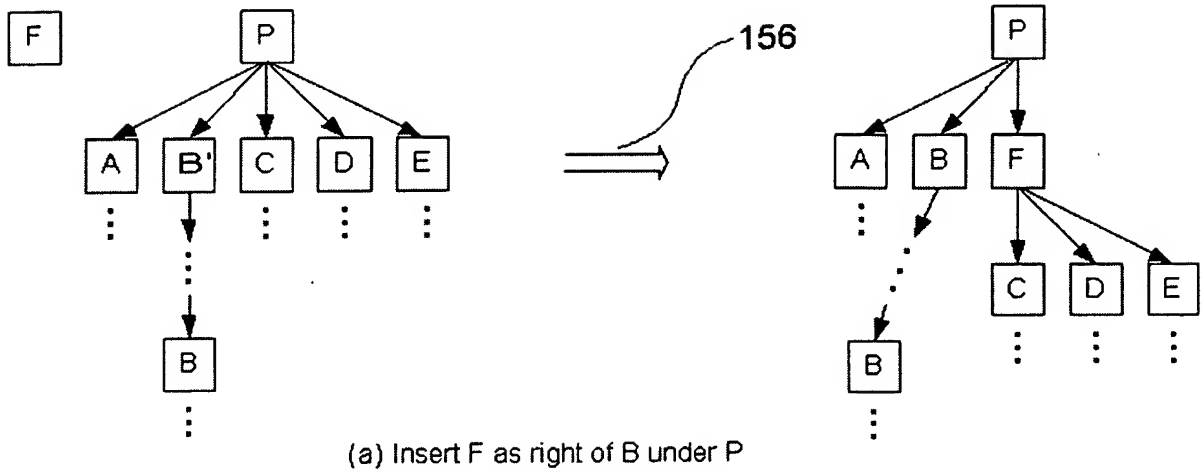
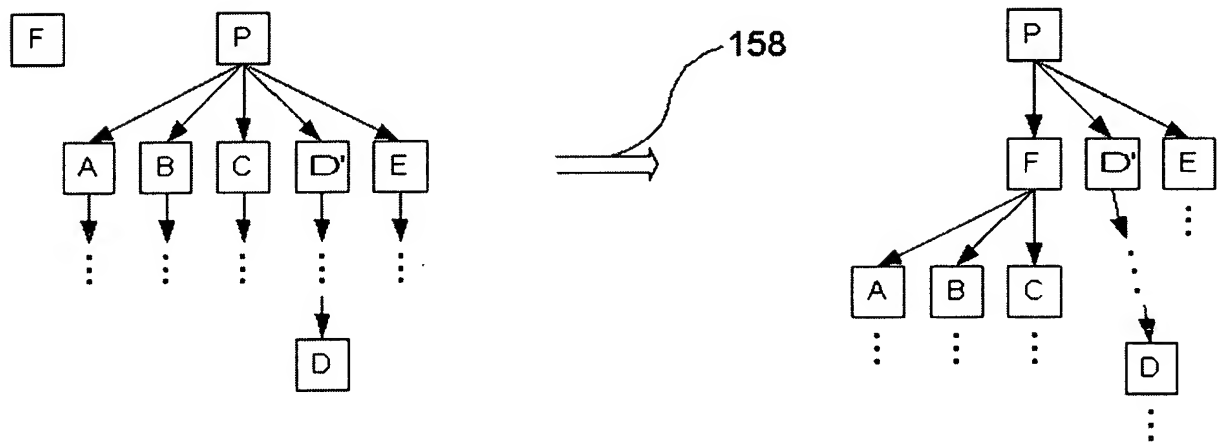


Figure 16

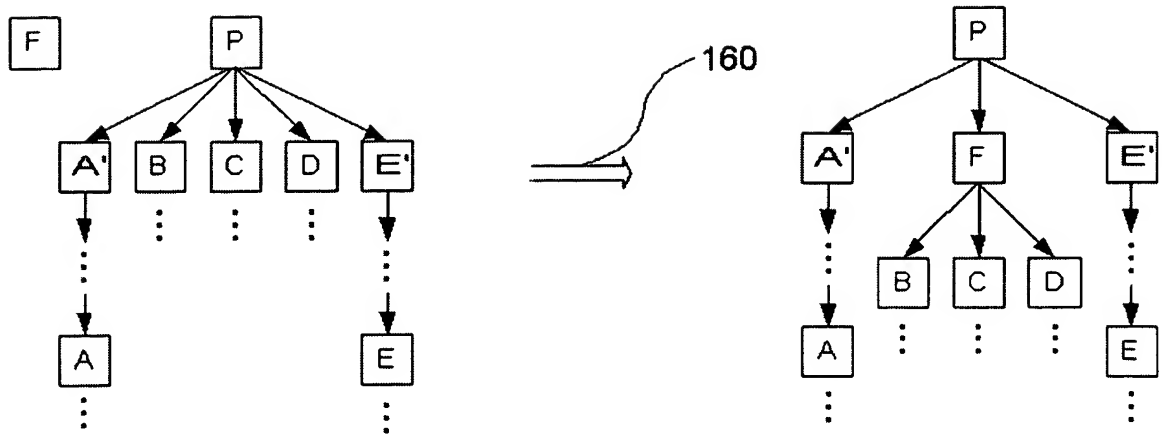


(a) Insert F as right of B under P

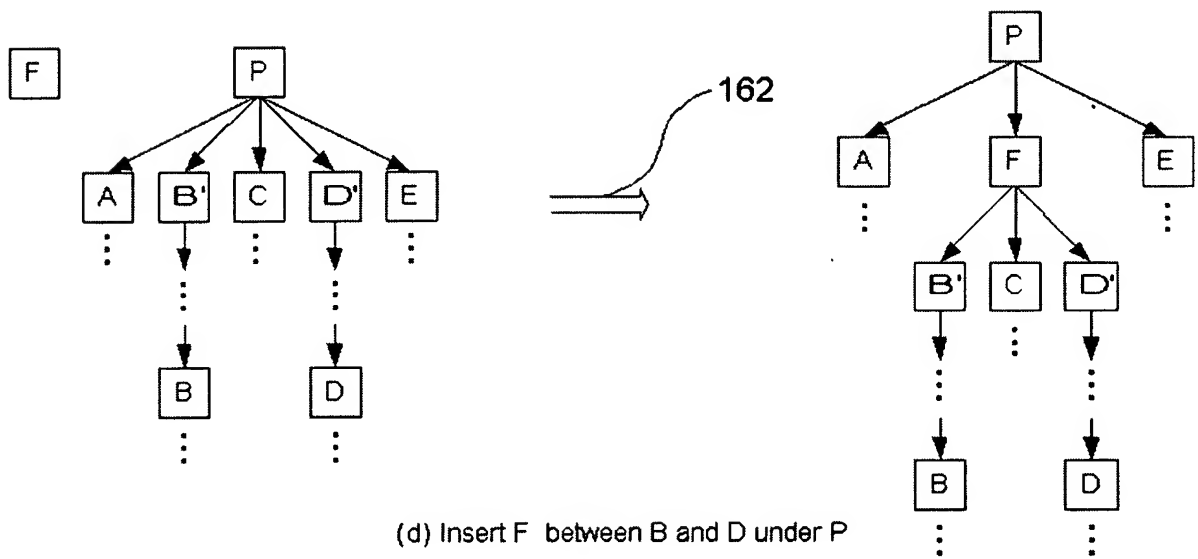


(b) Insert F as left of D under P

Figure 17a

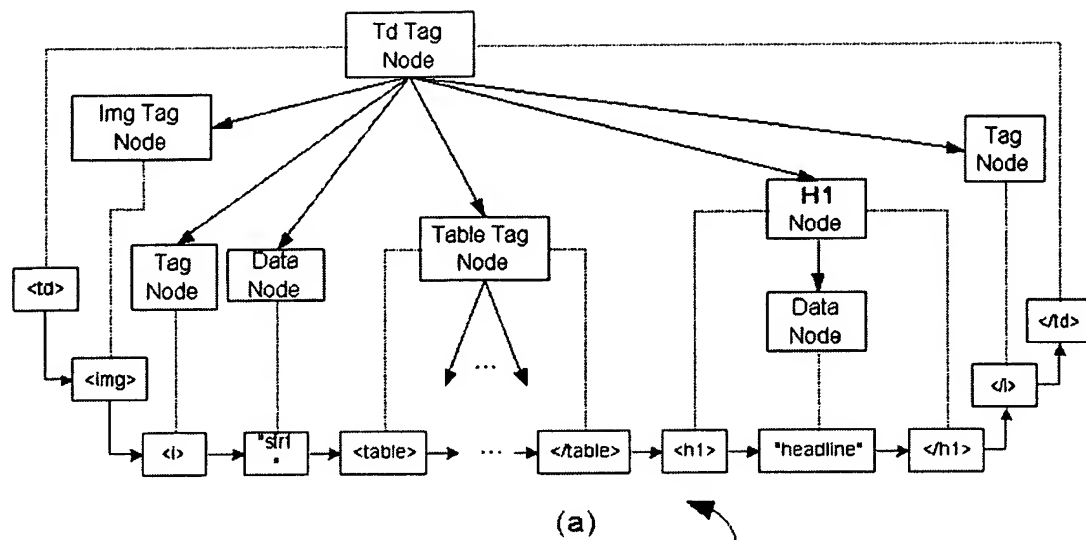


(c) Insert F as between A and E under P



(d) Insert F between B and D under P

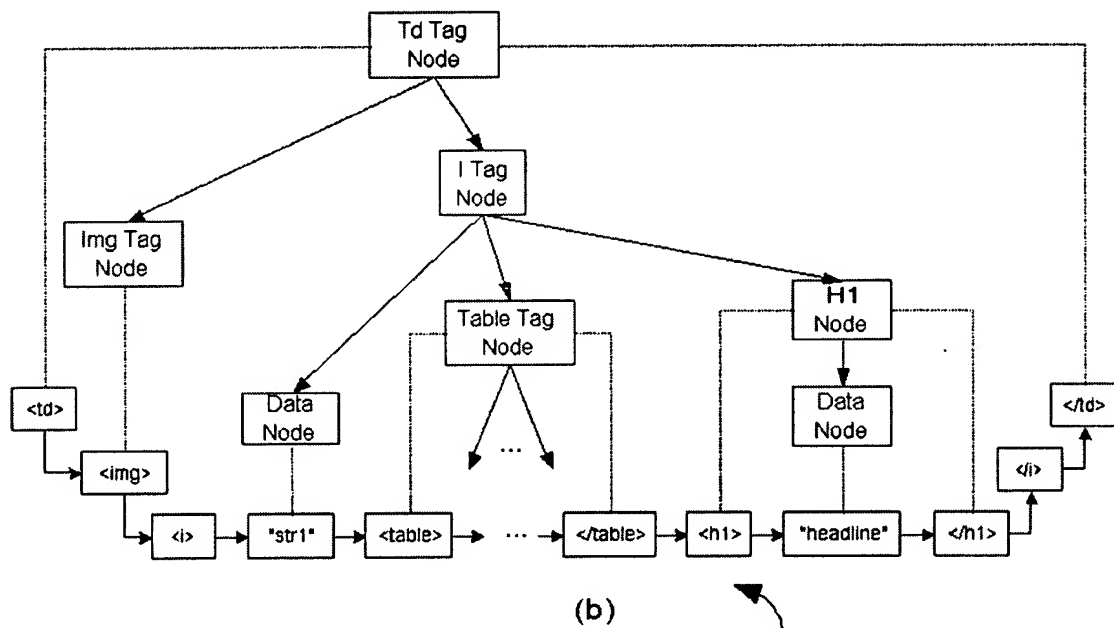
Figure 17b



147

<td><img><i>str1<table>...  
</table><h1>headline</h1></i></td>

147a



147b

Figure 18



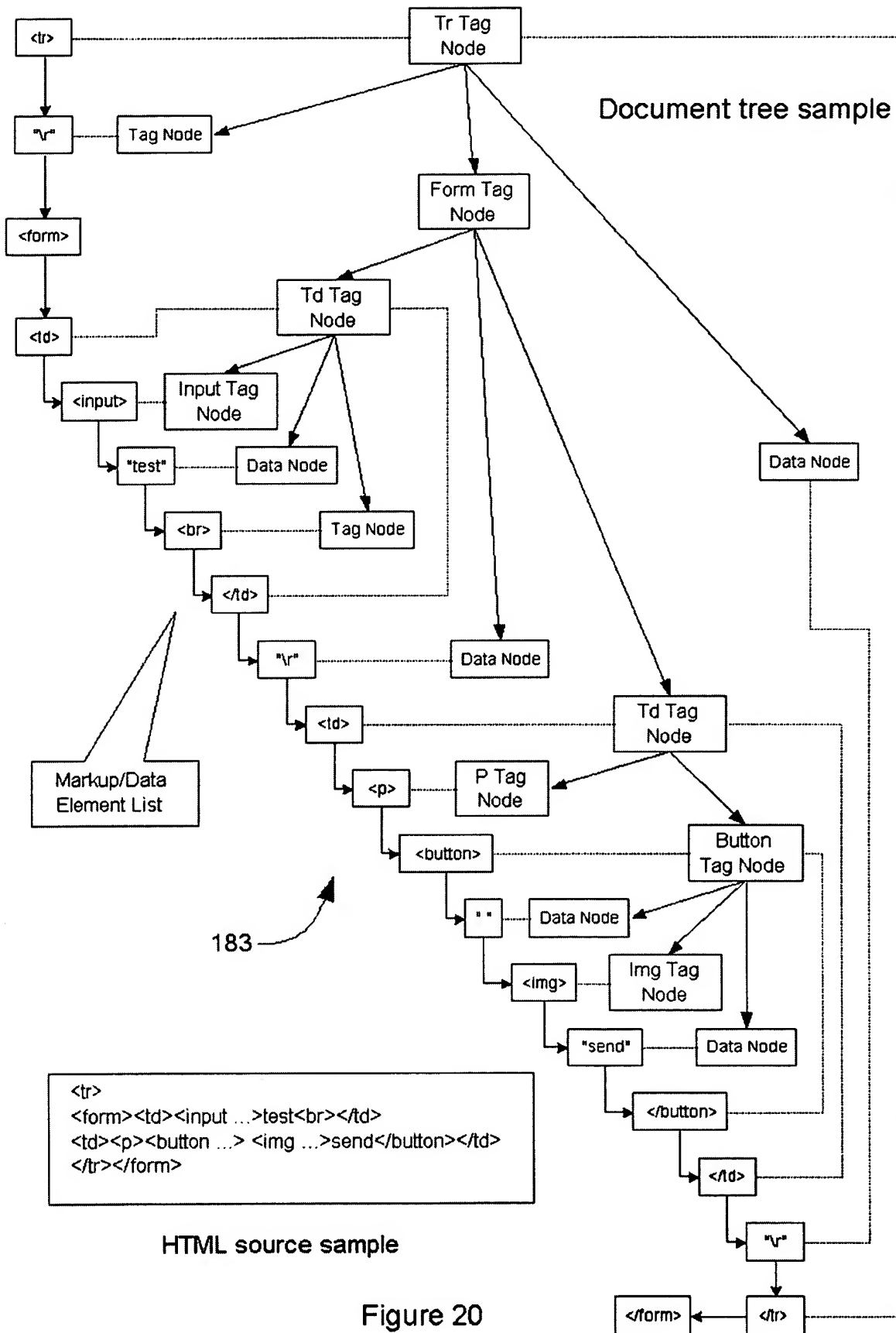


Figure 20

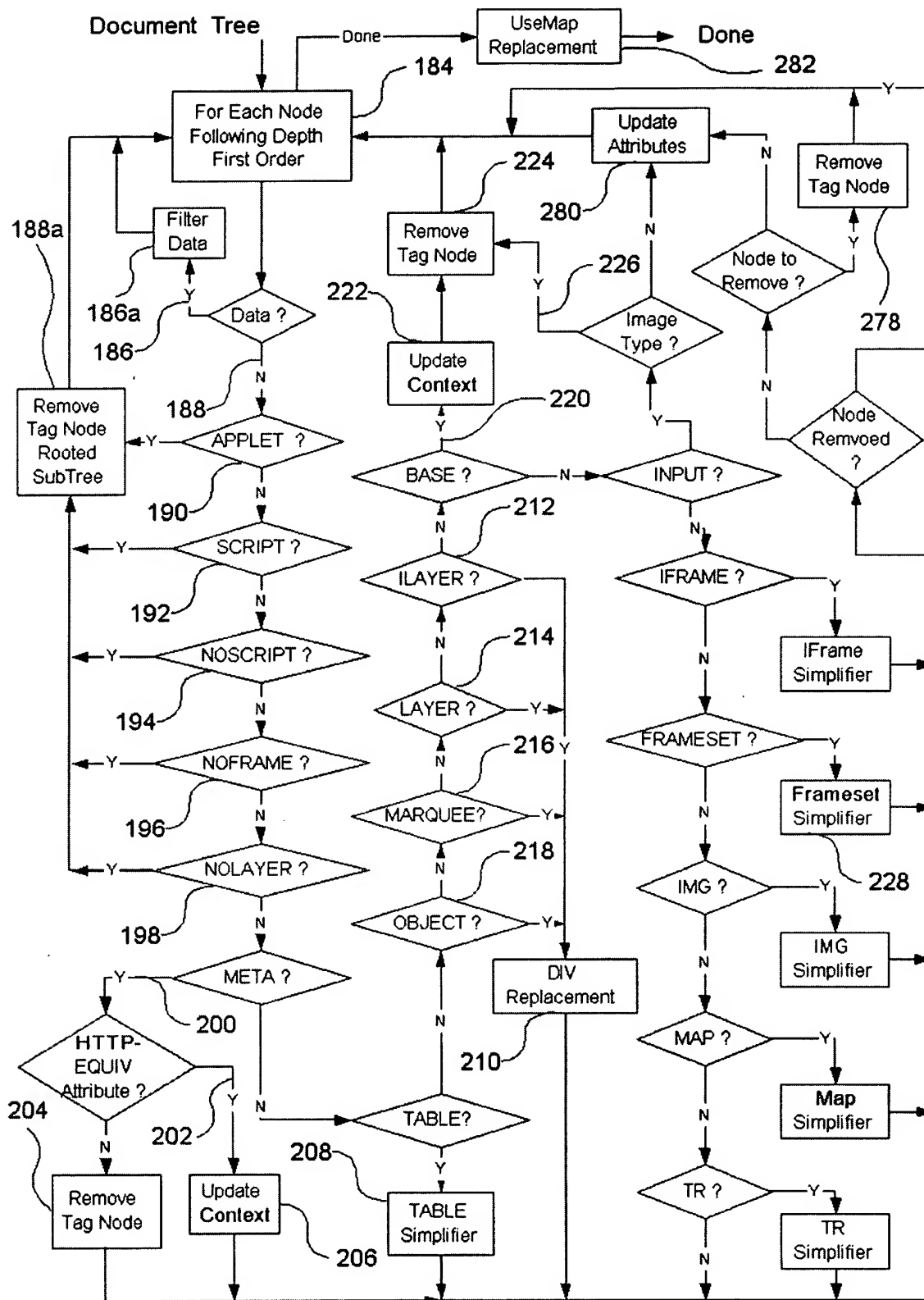
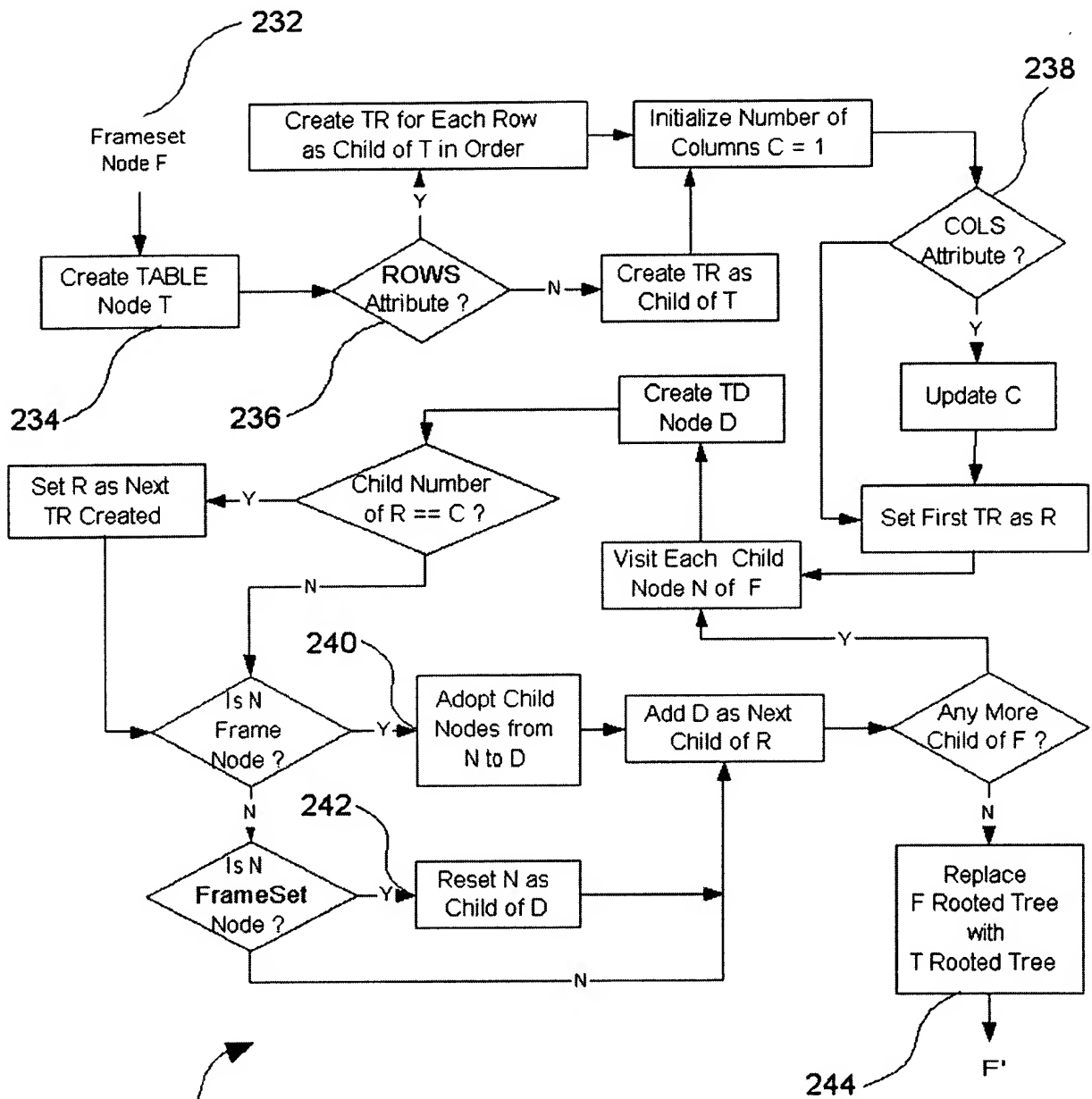


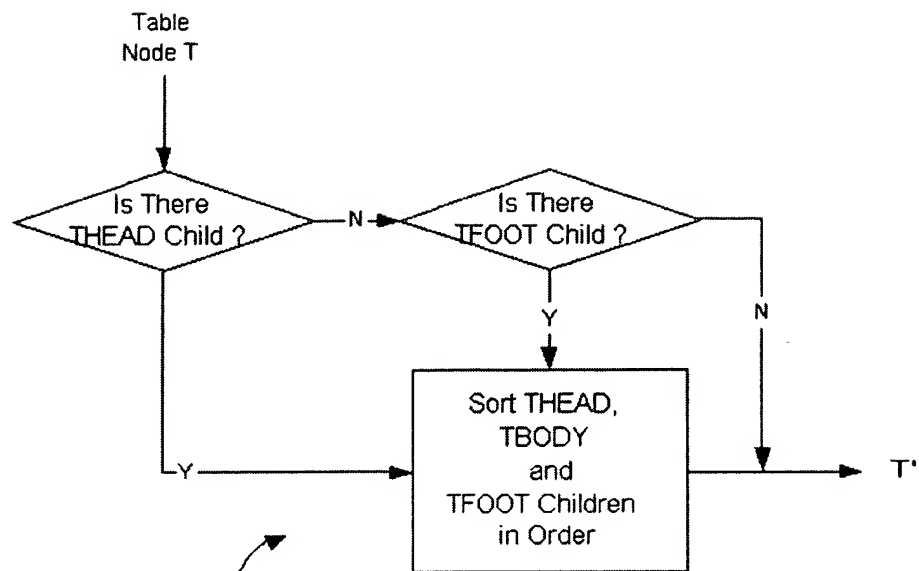
Figure 21



(a) Frameset Simplifier

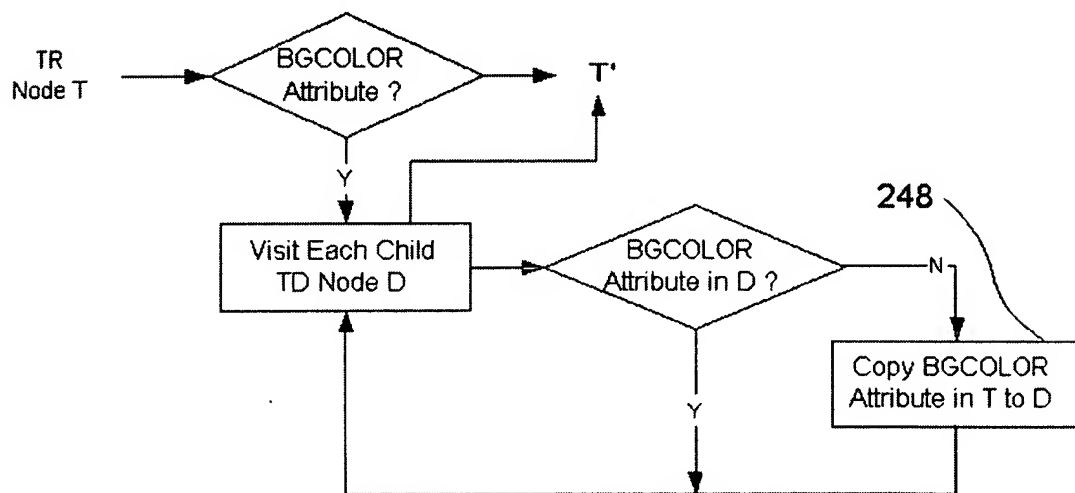
Figure 22a





209

(b) Table Simplifier



248

246

(c) TR Simplifier

Figure 22b

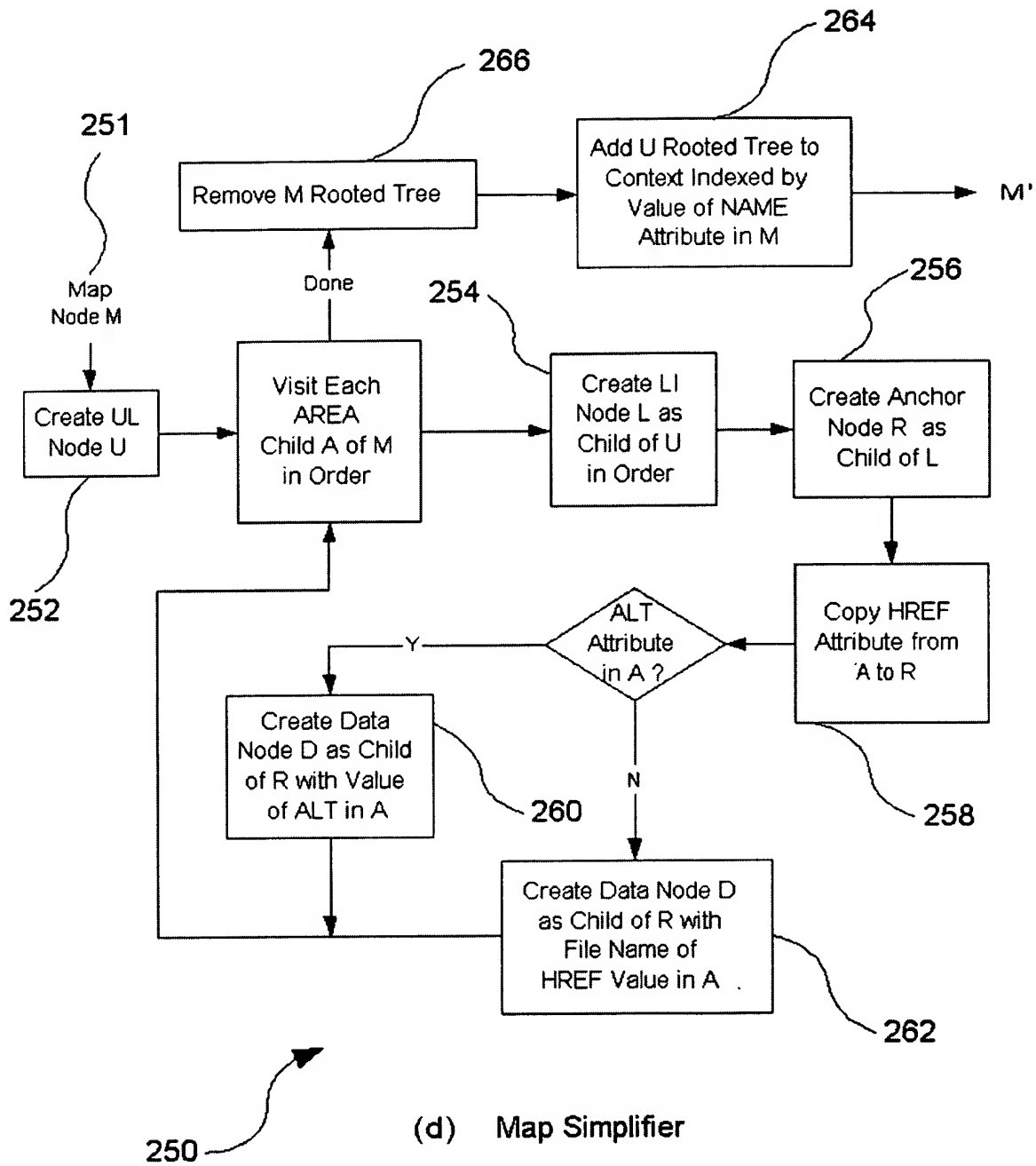
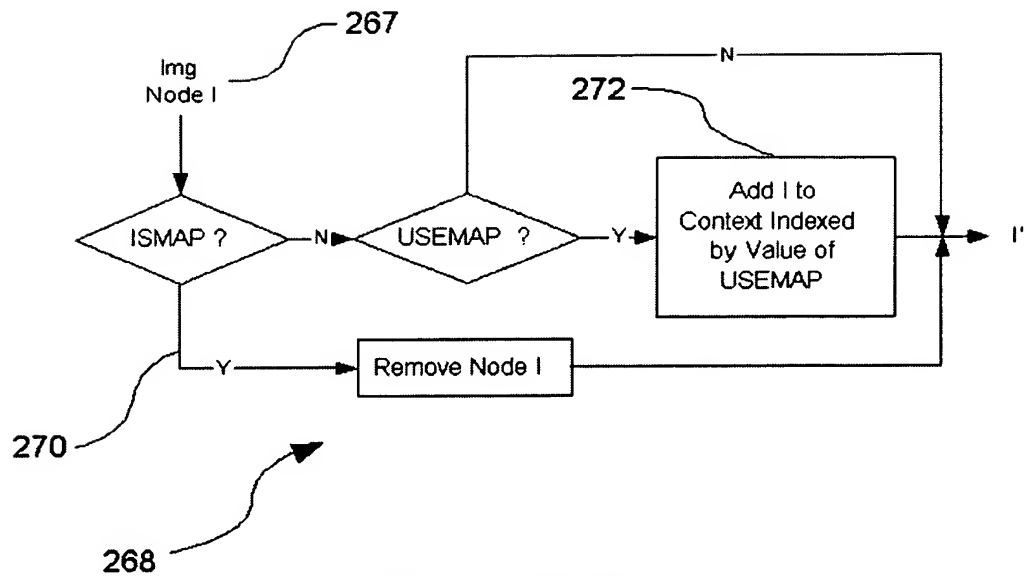
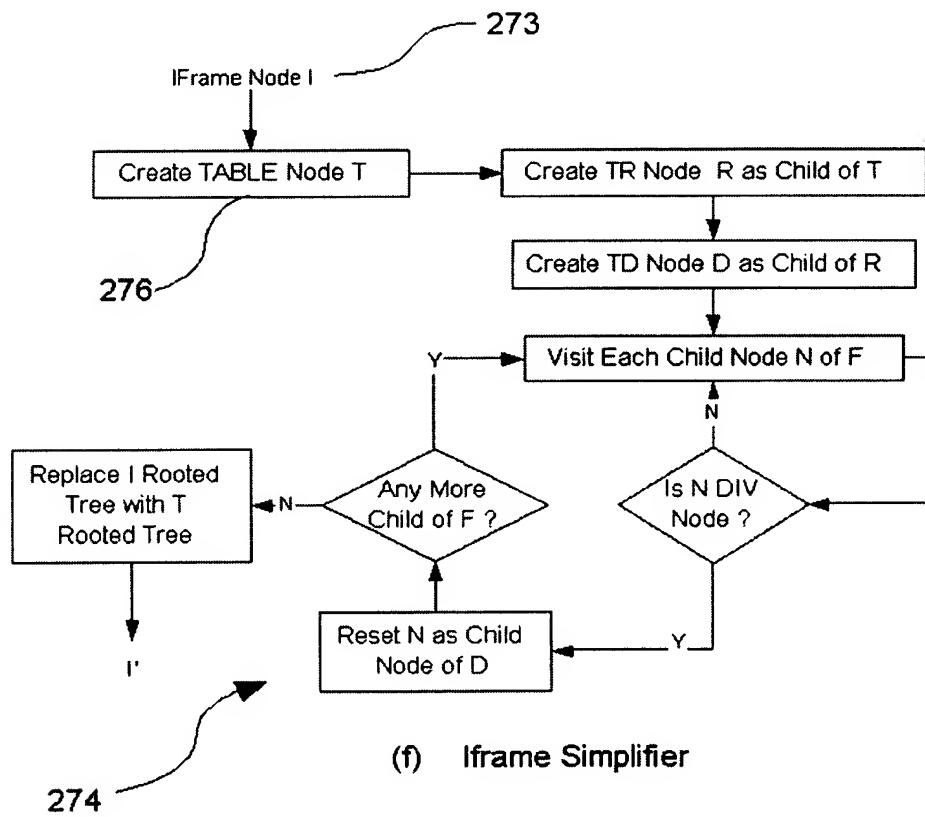


Figure 23a



(e) Img Simplifier



(f) Iframe Simplifier

Figure 23b

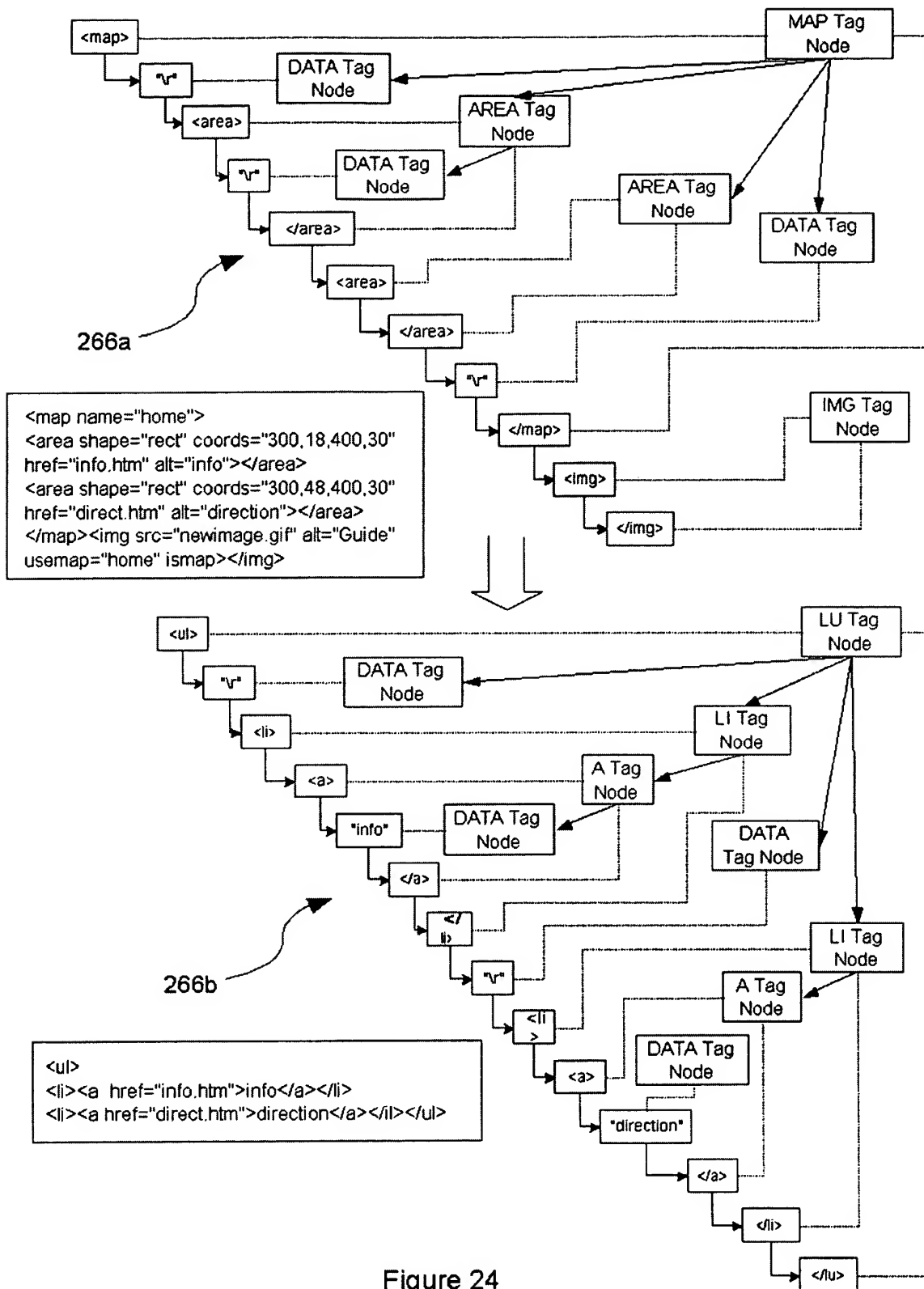


Figure 24

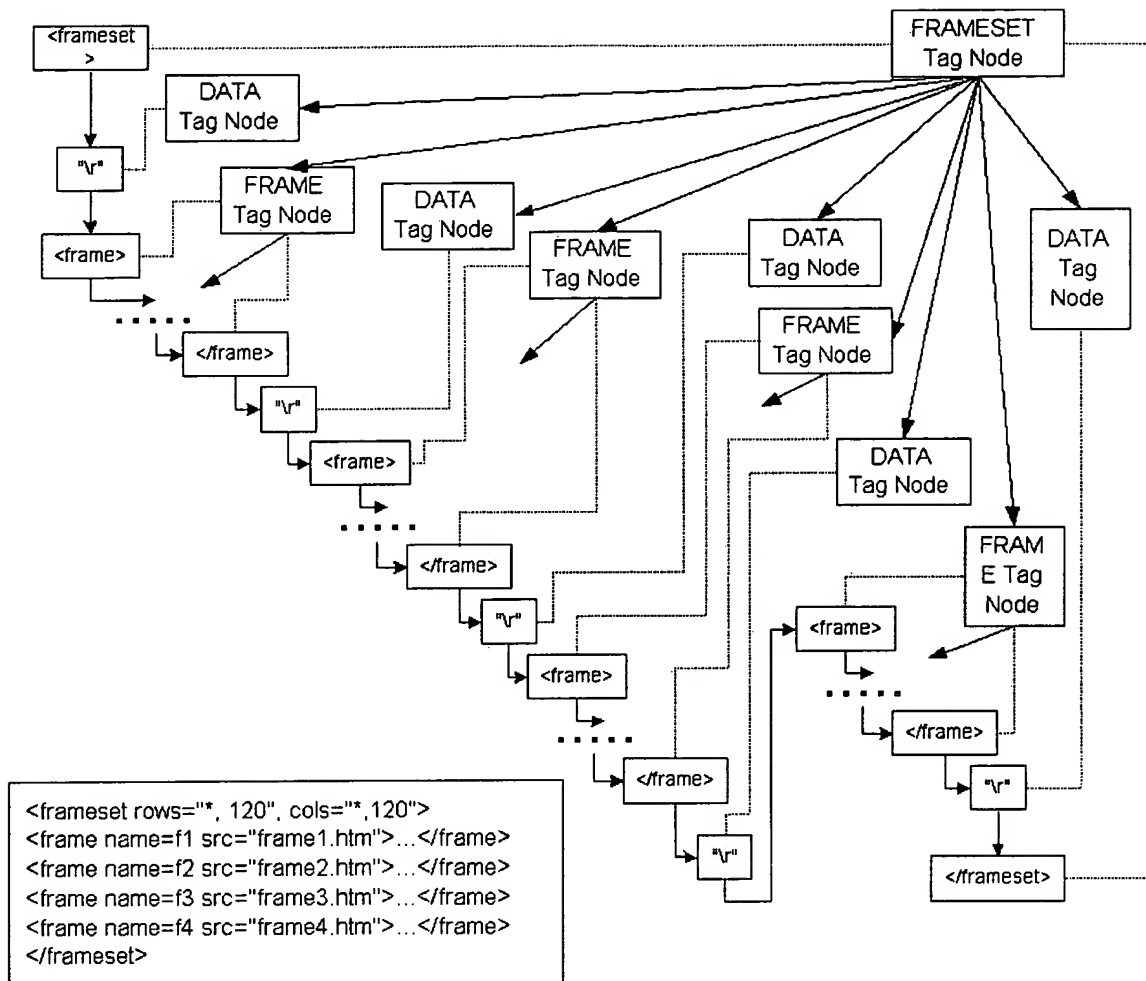


Figure 25



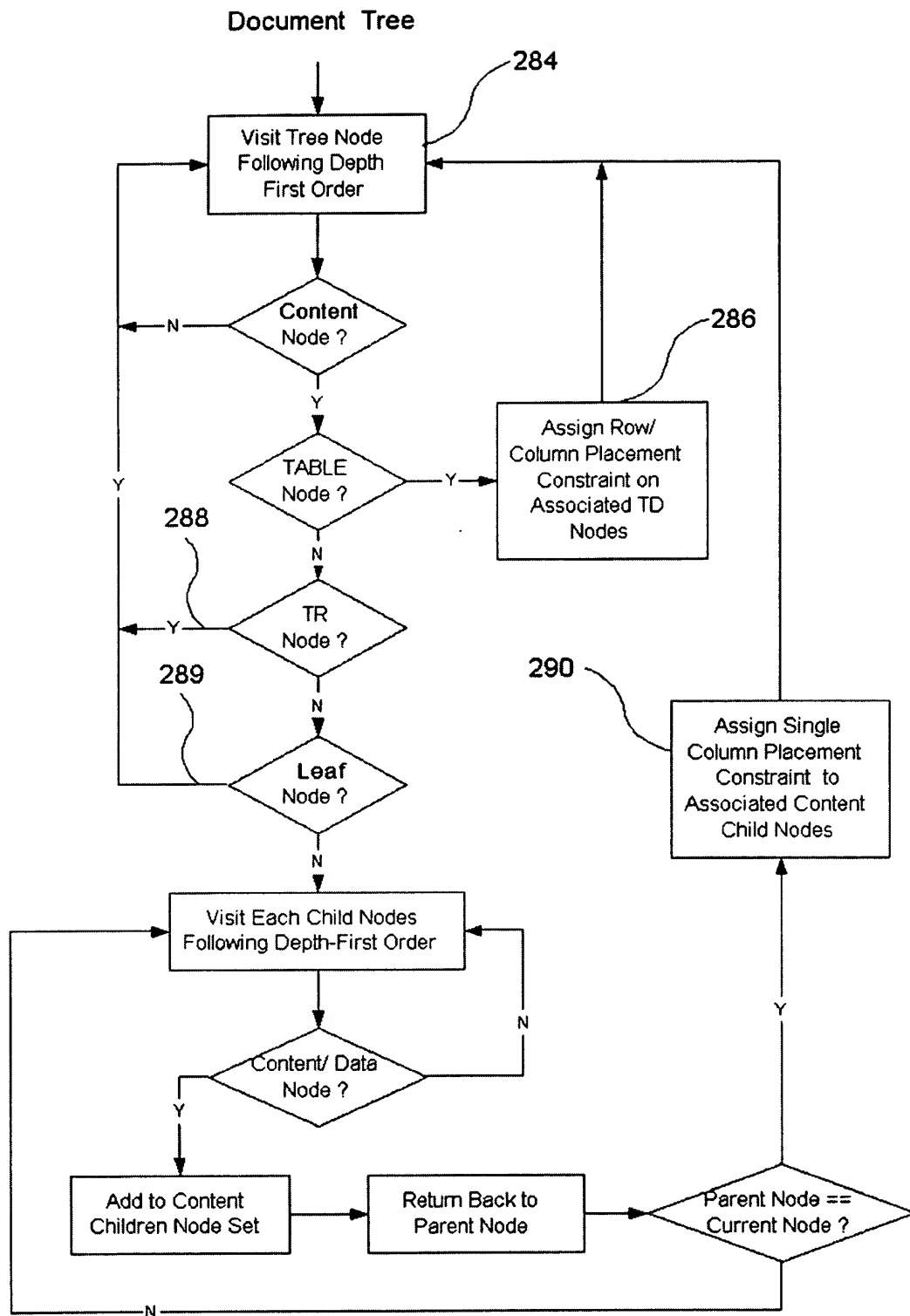


Figure 27

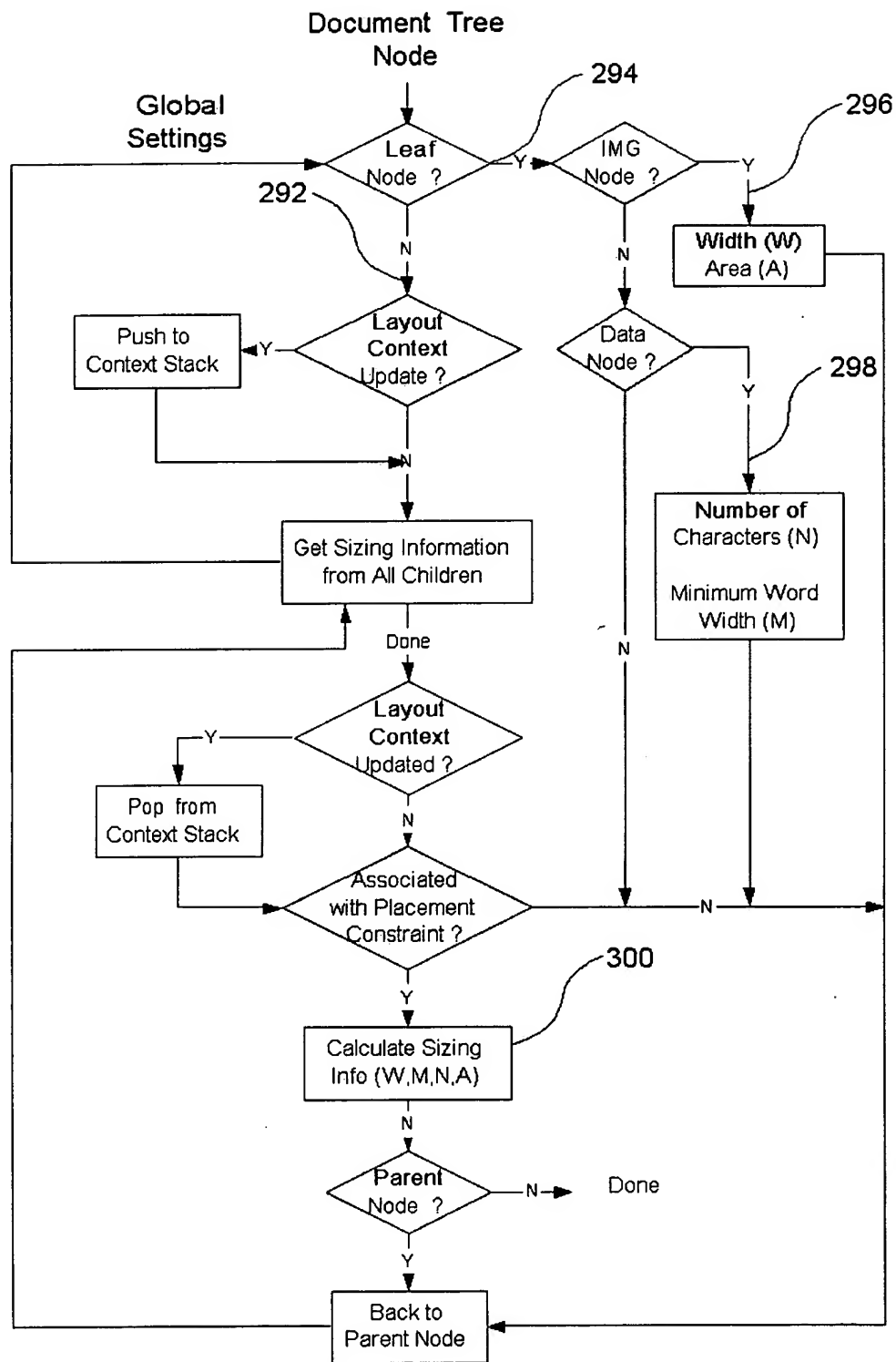


Figure 28



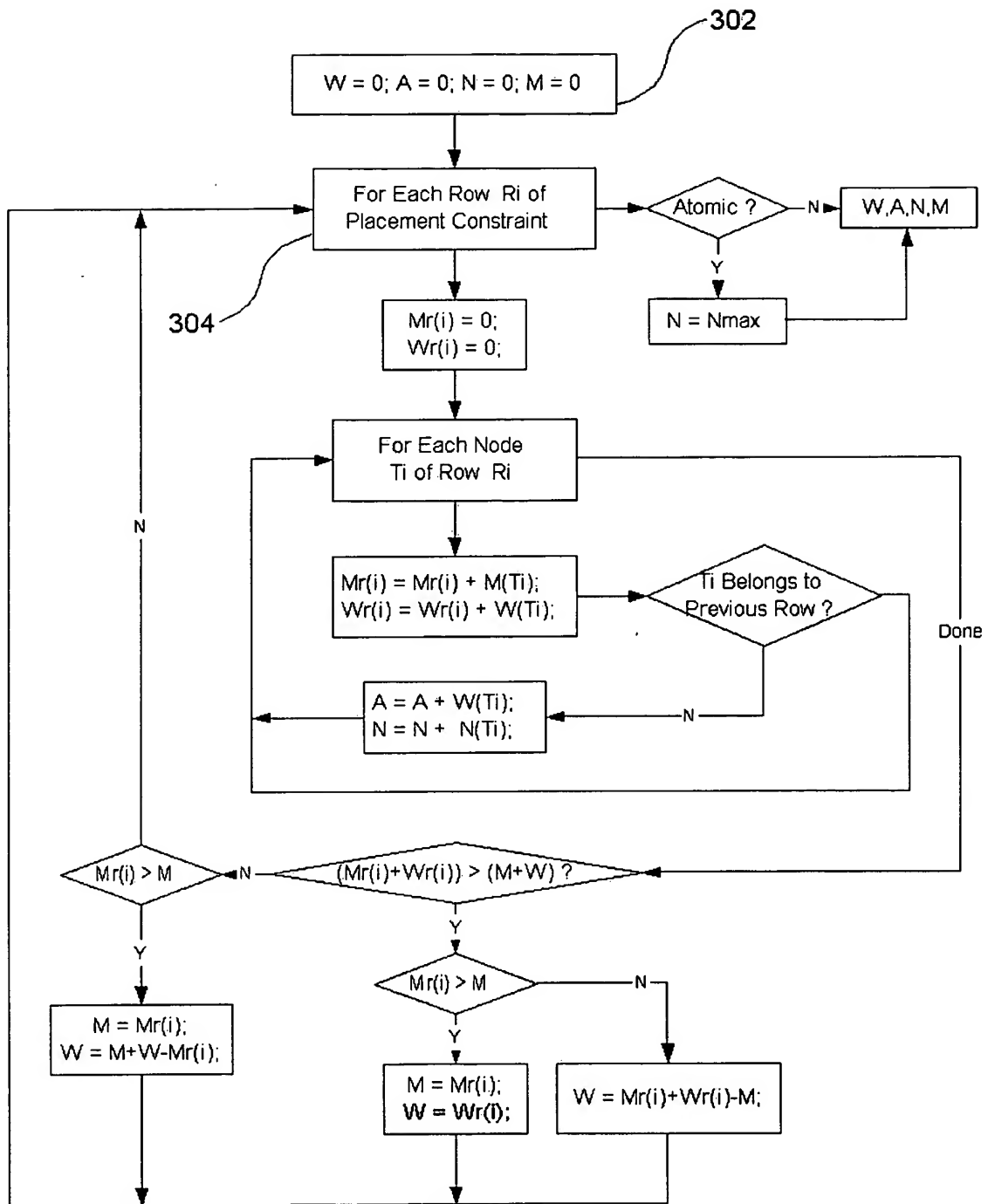


Figure 29

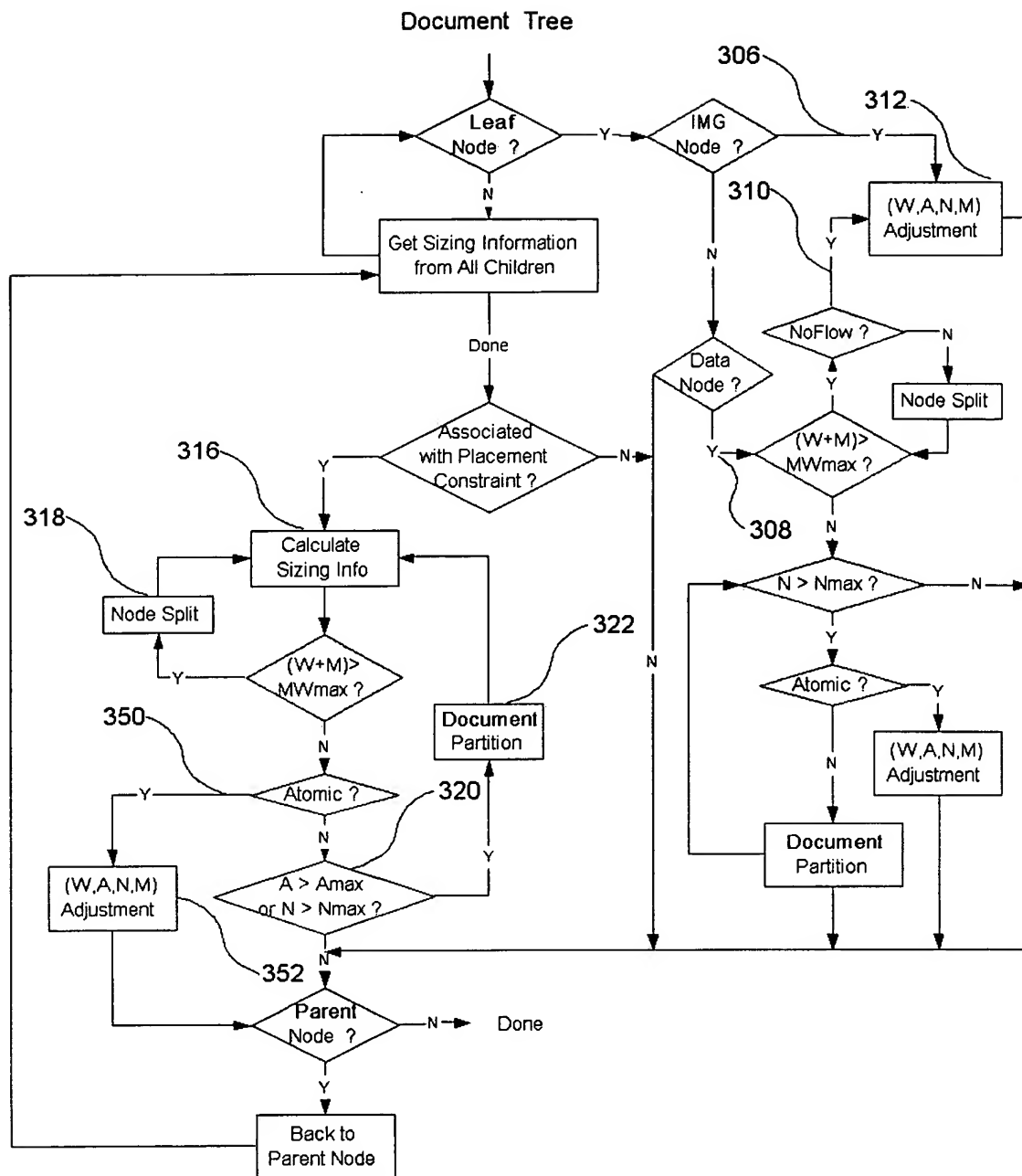


Figure 30

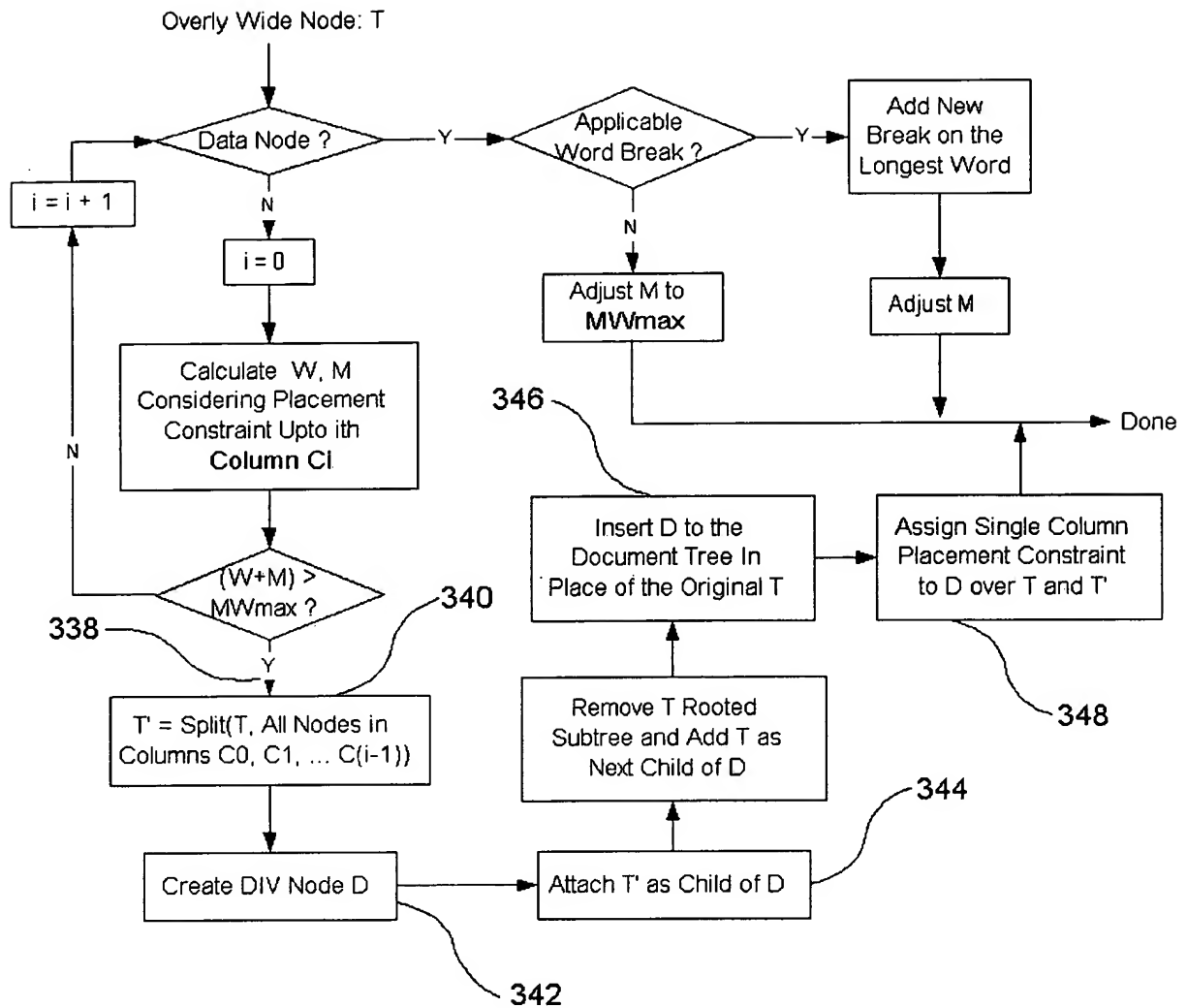


Figure 31

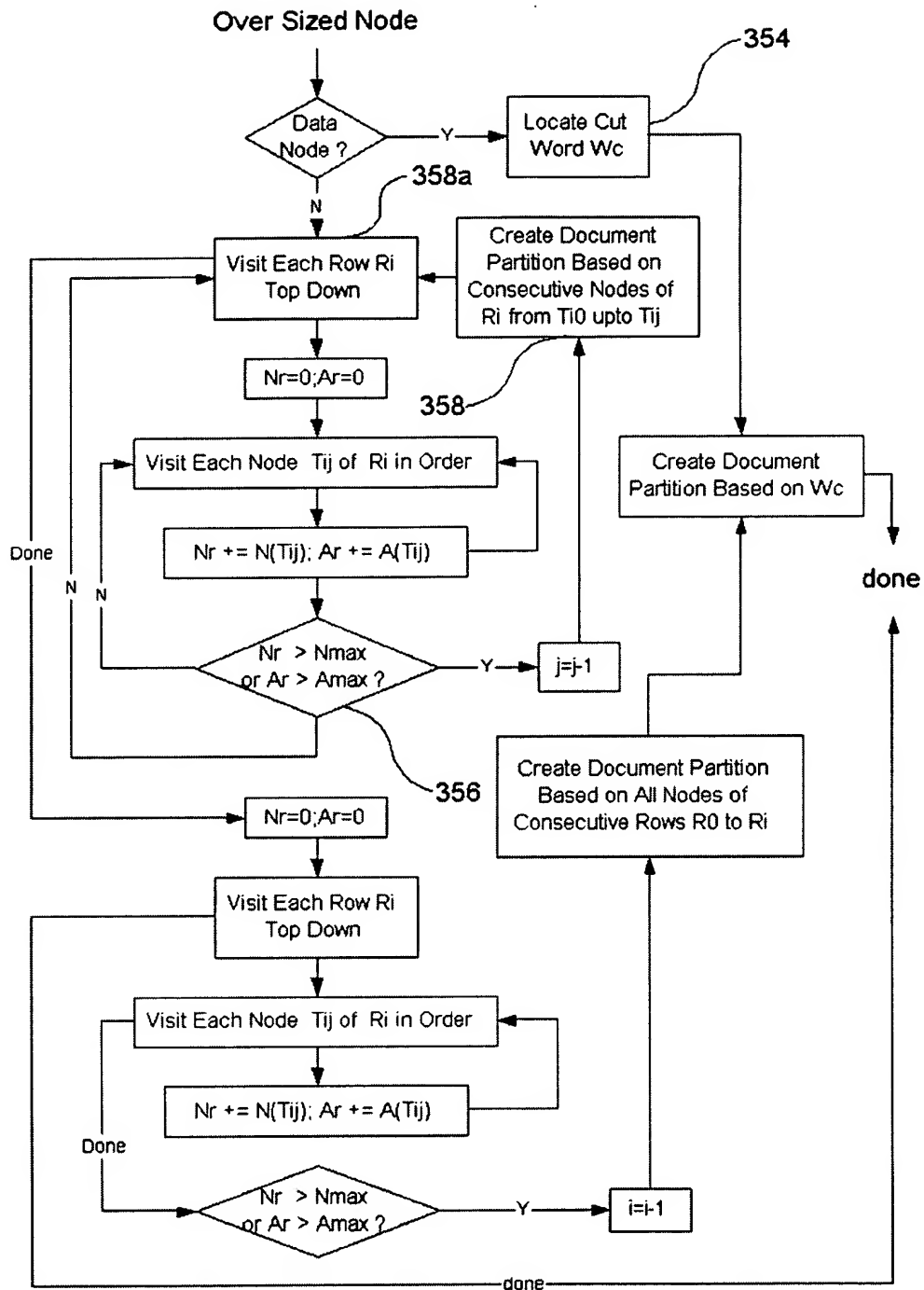
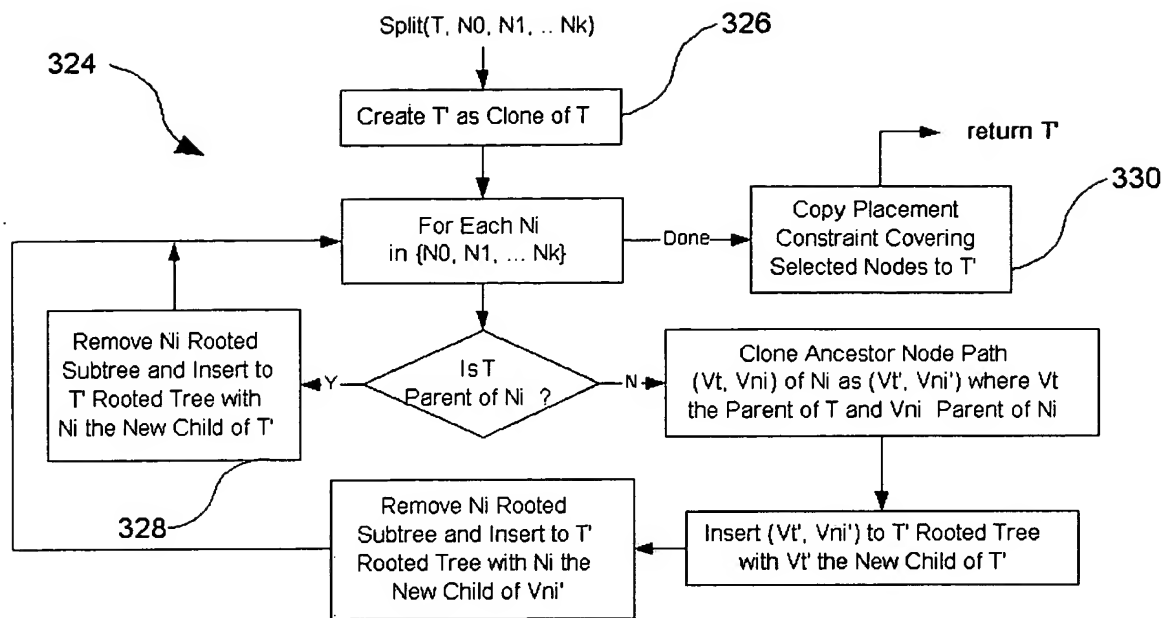


Figure 32



(a) Steps to Split a  $T$  Rooted Tree with Selected Content Child Nodes  $N_0, N_1, \dots, N_k$ : Split( $T, N_0, N_1, \dots, N_k$ )

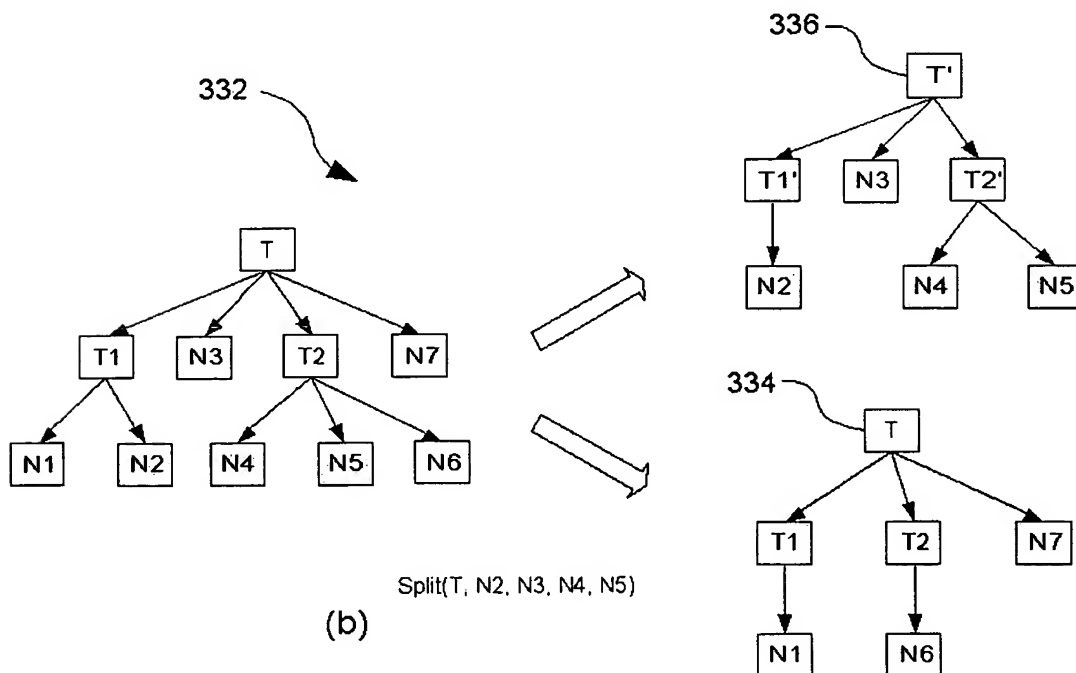


Figure 33

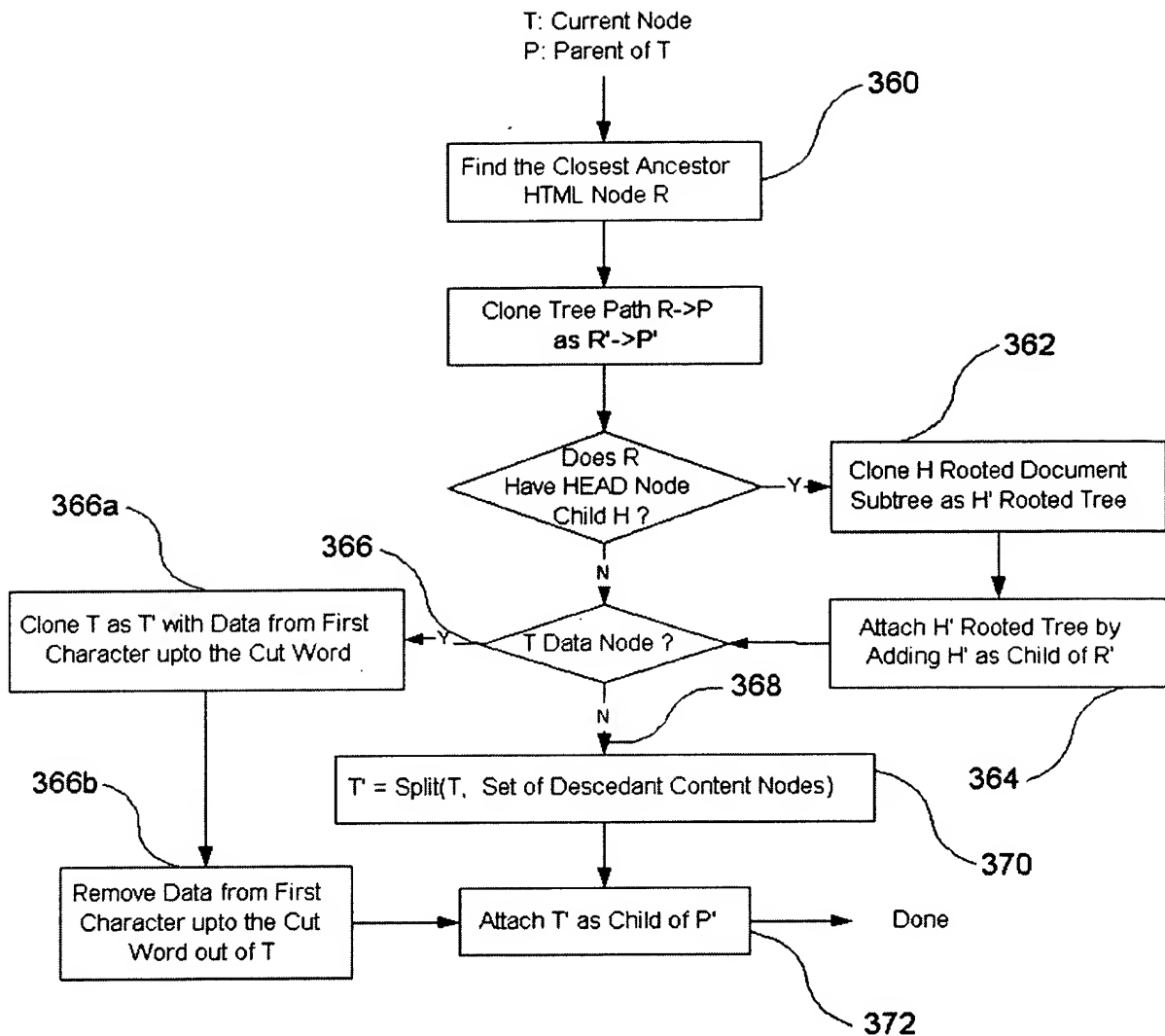


Figure 34

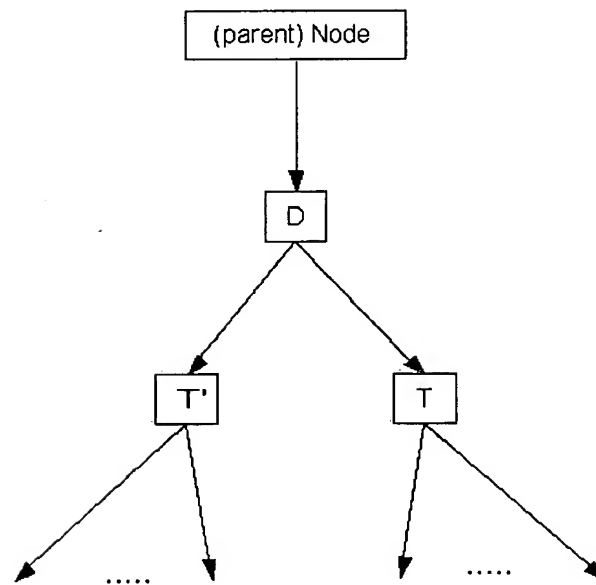
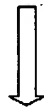
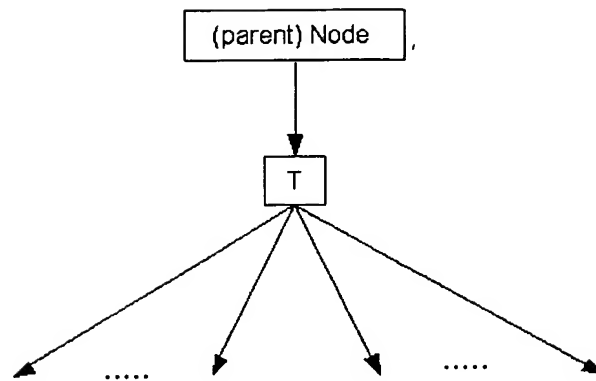


Figure 35

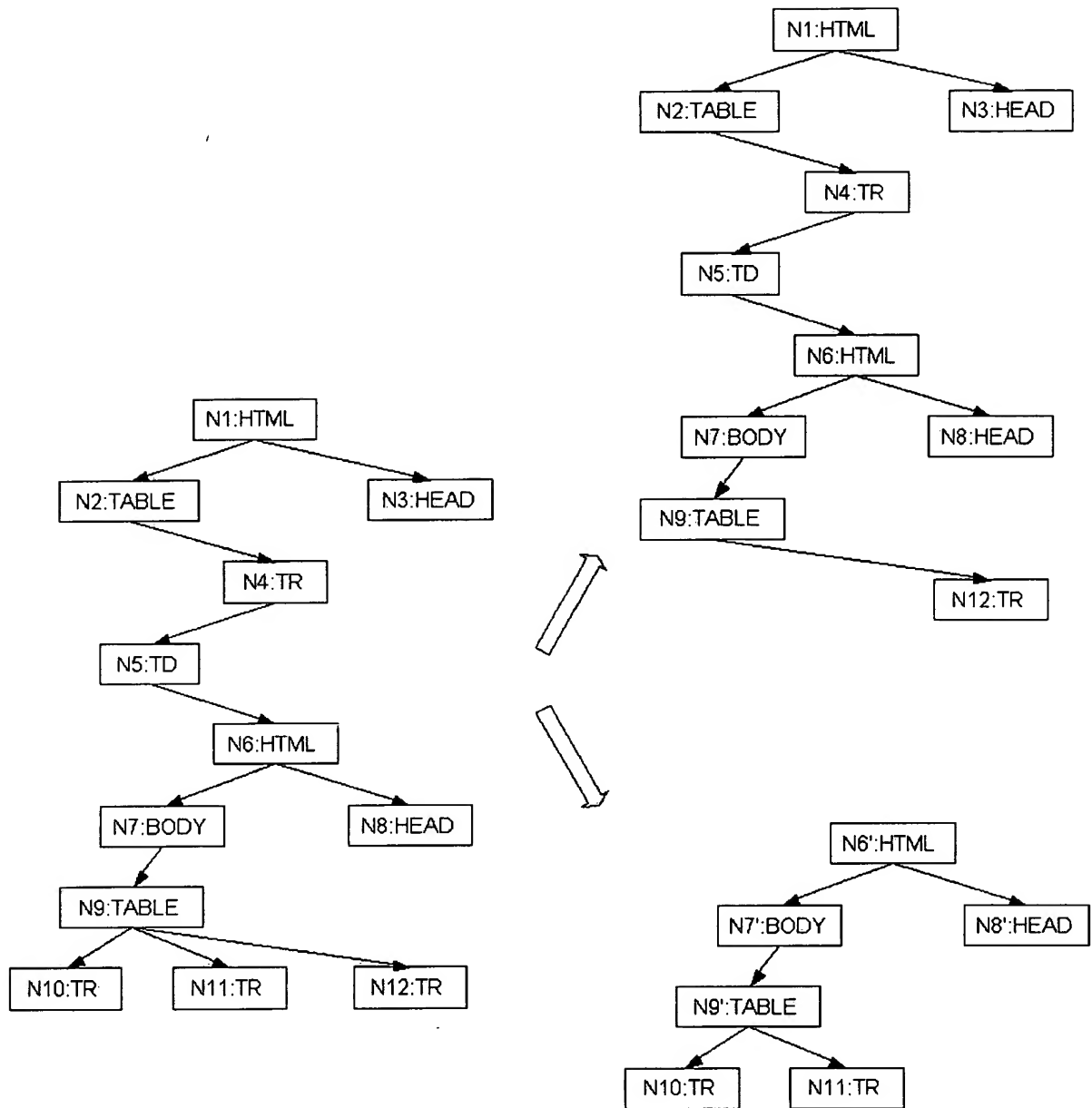


Figure 36



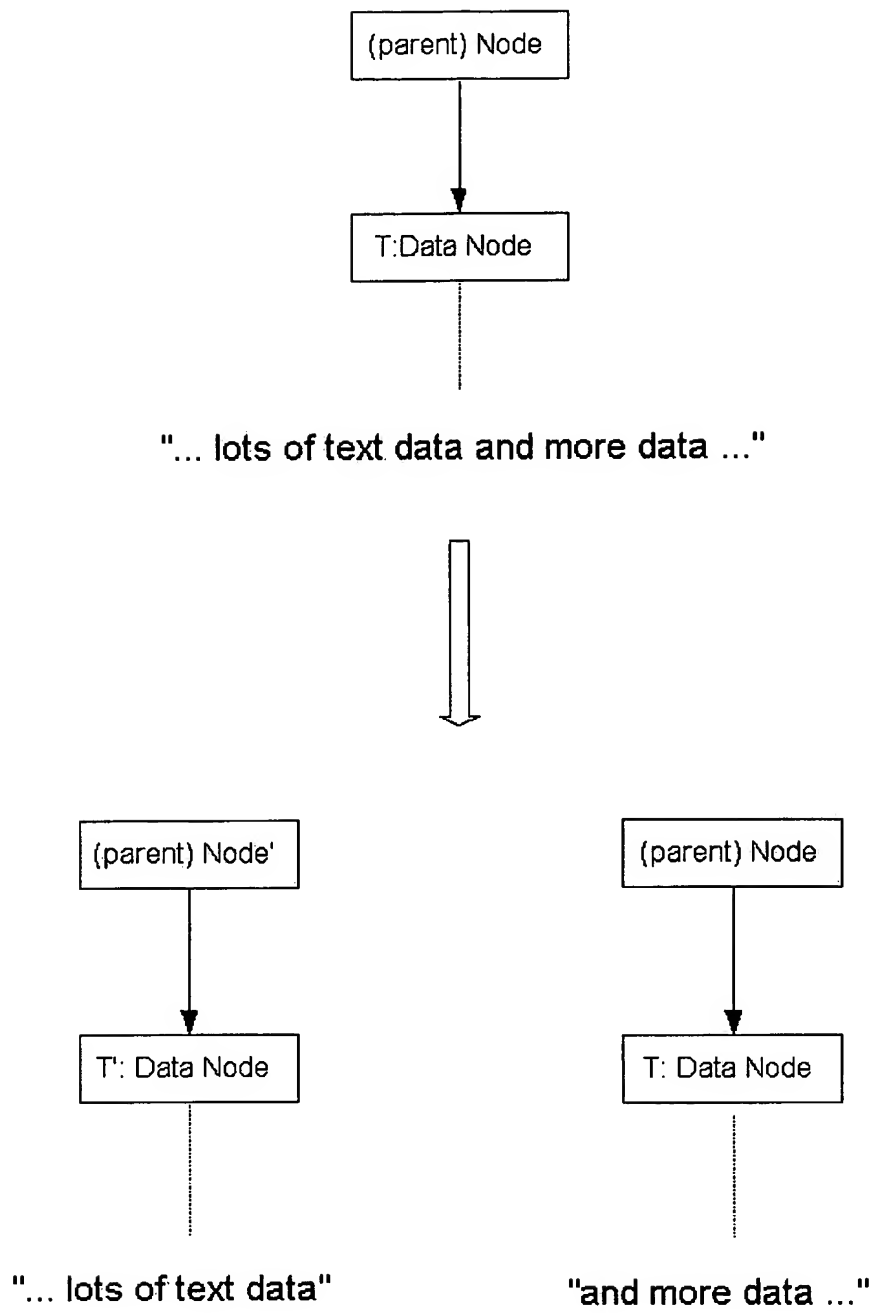


Figure 37

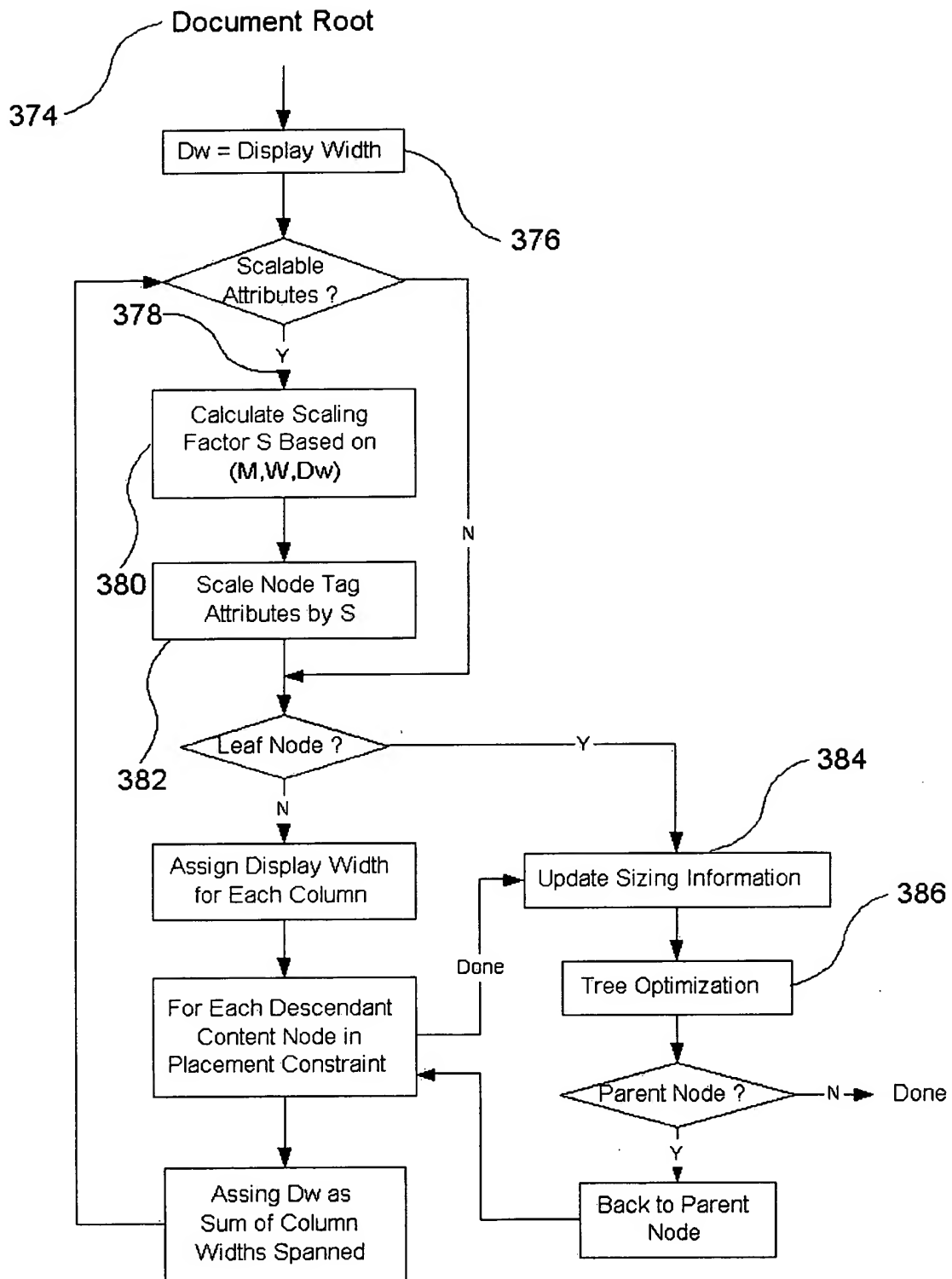


Figure 38

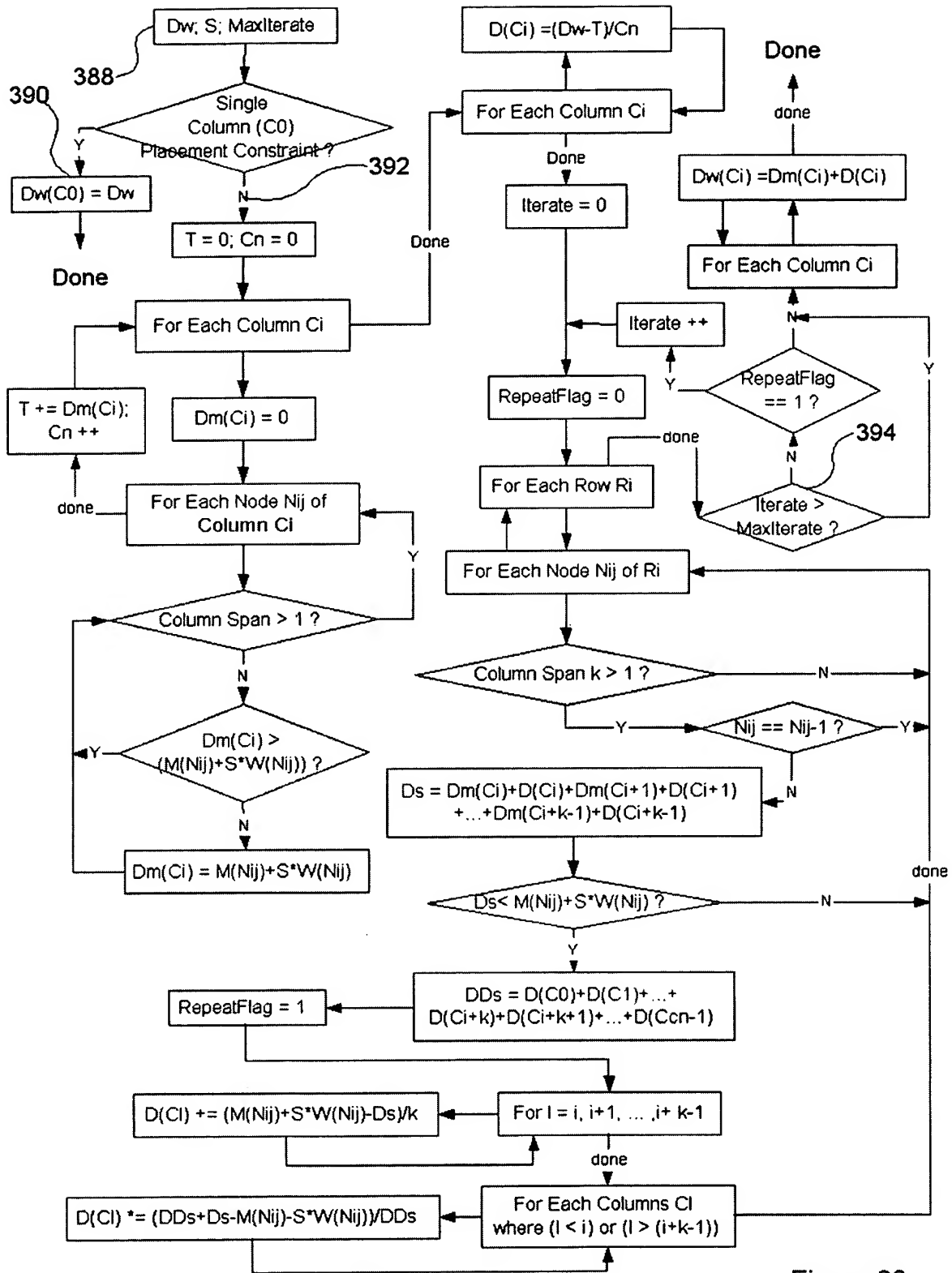


Figure 39

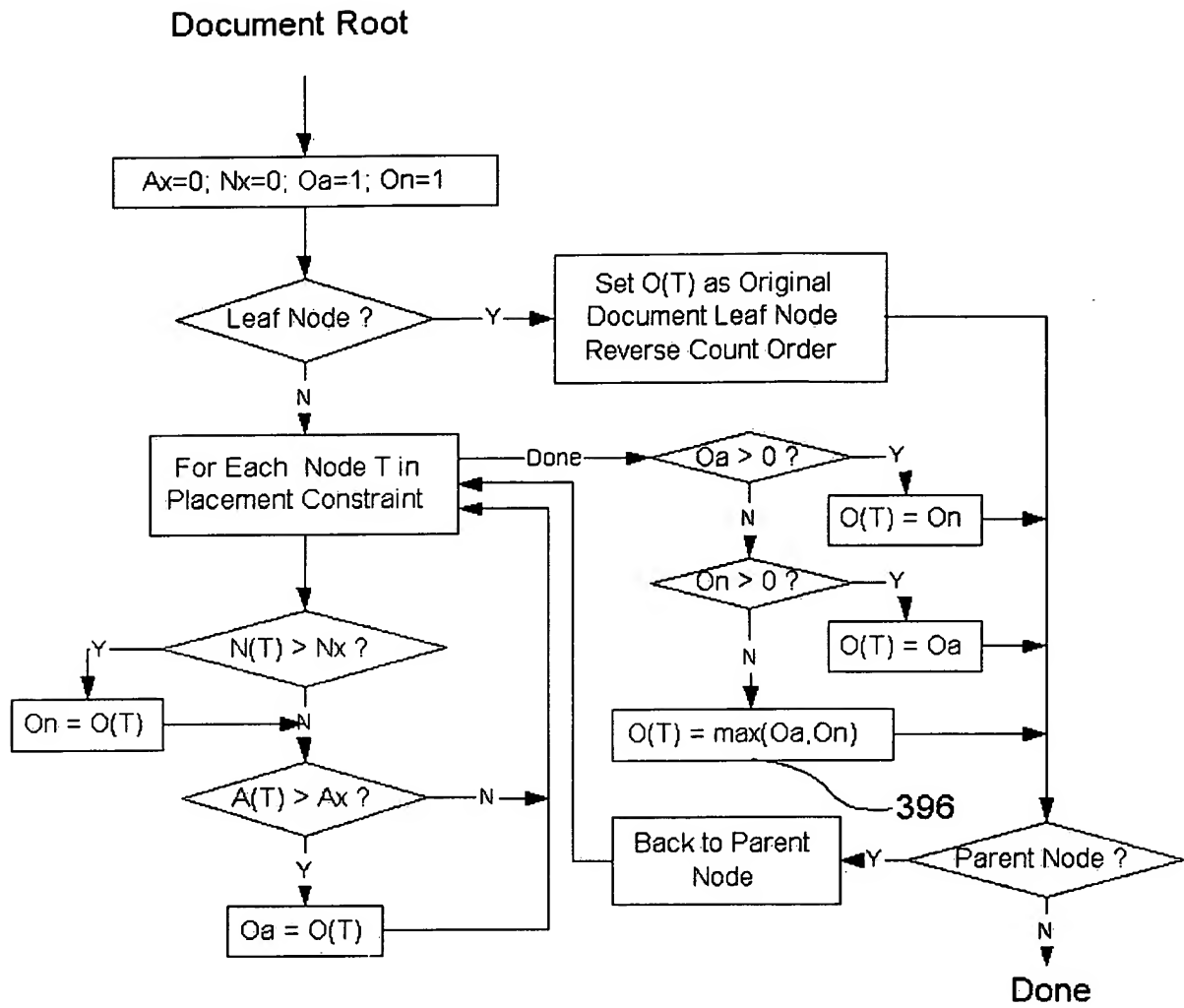


Figure 40

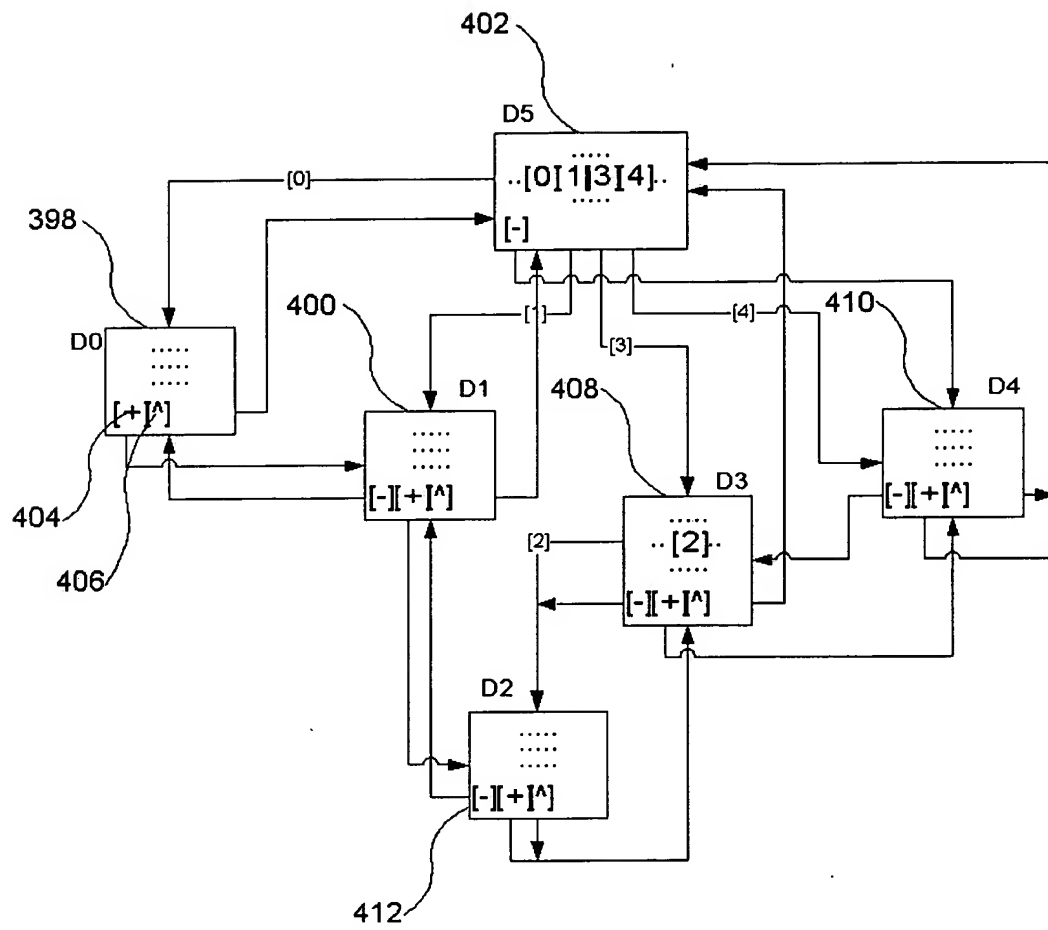


Figure 41